Stephen J. Joncus, OSB #013072
Email: stephen.joncus@klarquist.com
John D. Vandenberg, OSB #893755
Email: john.vandenberg@klarquist.com
KLARQUIST SPARKMAN, LLP
121 S.W. Salmon Street, Suite 1600
Portland, Oregon 97204
Telephone: 503-595-5300
Facsimile: 503-595-5301

Attorneys for Plaintiff Digimarc Corporation

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF OREGON
PORTLAND DIVISION

CV'09 1355 KI

Case No.: _____

Digimarc Corporation,

Plaintiff,

vs.

Shazam Entertainment, Ltd.,

Defendant

COMPLAINT FOR PATENT INFRINGEMENT

DEMAND FOR JURY TRIAL

#30132

Plaintiff Digimarc Corporation ("Digimarc") makes the following allegations for its Complaint of Patent Infringement against defendant Shazam Entertainment, Ltd. ("Shazam"):

ASSERTED PATENTS

1. This is a patent infringement case involving U.S. Patent Nos. 7,349,552 ("the '552 patent"), 7,587,602 ("the '602 patent"), and 7,590,259 ("the '259 patent"). The '552 patent issued March 25, 2008; the '602 patent issued September 8, 2009; and the '259 patent issued September 15, 2009. The '552 patent and '259 patent each issued from a series of applications extending back to July 27, 1995. The '602 patent issued from applications extending back to May 19, 1999. Each of these patents issued to Plaintiff or its predecessor, and is assigned to and owned by Plaintiff Digimarc Corporation.

BACKGROUND

2. Digimarc is an industry leader in content identification technologies. Digimarc's technology, and those of its licensees, has been broadly deployed across a range of applications spanning audio, video, images, and printed materials. For example, essentially all of the broadcast television in the United States is identified using Digimarc's technology. Similarly, most of the driver licenses issued in the U.S. since 2007 incorporate Digimarc's technology. With a technical staff of fifty (several with PhDs), the company actively partners with many other companies in developing leading edge market-focused solutions – providing them the benefit of Digimarc's scientific and industry experience.

3. In today's media rich world, consumers are exposed to music in many different ways, including through movies and TV shows, night clubs, parties, etc. Identifying and then following up and learning more about that music is as difficult as it ever was. At the same time, Internet technologies have built an expectation among consumers that immediate information can be obtained on a whim, preferably via the cell phones that many carry everywhere. Today's music fans, like those in the past, experience an emotional connection with a new song, yet now, they expect to be able to easily and immediately identify the song, learn more about it, and

conveniently acquire the song online. Geoff Rhoads, the founder of Digimarc, foresaw the need for technology to enable devices to identify audio and immediately link the consumer to associated Internet services. His 1995 and 1996 patent applications are foundations on which popular music identification services like Shazam have been built.

4. Shazam Entertainment Ltd. offers a music identification service, called Shazam, using the content identification and web linking technology invented by Mr. Rhoads. In one popular implementation, Shazam's software is used by a consumer to record a brief song excerpt using the microphone of a cell phone. This excerpt is then analyzed to derive identifying data. Shazam then uses the Internet to send information to the consumer's phone identifying the song, and presenting related information. Among the related information commonly provided are musician biography and "discography" information, and opportunities for the consumer to purchase the song from an online vendor and view online concert videos.

5. The U.S. Patent Office issued Digimarc the '552, '602, and '259 patents for some of its inventions on this technology.

6. Shazam's Shazam application and service infringes at least these three Digimarc patents.

7. Upon information and belief, Shazam's application and service has been used by citizens of Oregon, in the state Oregon.

8. Digimarc wrote Shazam in June, 2009, alerting Shazam to six of Digimarc's patents - particularly explaining the application of certain claims from those patents to Shazam's application and service offerings, and inviting Shazam to engage in discussions concerning a resolution that could include a collaborative business relationship exploiting shared interests of the companies. Despite repeated communications from Digimarc since June, Shazam has failed to substantively engage with Digimarc to resolve the patent issues or stop its infringing actions. Digimarc, therefore, is asking this Court for relief for the infringing acts by Shazam.

THE PARTIES

9. Digimarc has its principal place of business in this judicial district at 9405 SW Gemini Drive, Beaverton, OR 97008.

10. On information and belief, Shazam is a registered corporation organized and existing under the laws of England, having a principal place of business at Charles House 4th Floor, 315 Kensington Street, London, England, W14 8QH, and has a registered office for purposes of service of process in Palo Alto, California.

JURISDICTION AND VENUE

11. This Court has subject matter jurisdiction under 28 U.S.C. §§ 1331 and 1338(a).

12. Shazam is subject to the personal jurisdiction of this Court. Shazam has sold and supplied its infringing applications and services in this judicial district for use in this judicial district, said use constituting ongoing infringement in this judicial district.

13. Shazam maintains interactive systems through which Oregon residents can use, and on information and belief at least thousands of Oregon residents have used, the infringing Shazam applications and services in this judicial district.

14. On information and belief, Shazam knows and intends that at least thousands of Oregon residents obtain, purchase and use its Shazam applications and services within this judicial district.

COUNT ONE – PATENT INFRINGEMENT

United States Patent No. 7,349,552

15. Digimarc incorporates the allegations stated in Paragraphs 1-14 as if fully set forth herein.

16. Digimarc is the sole owner of the entire right, title, and interest in United States Patent No. 7,349,552. A true copy of the '552 patent is attached as Exhibit A.

17. Digimarc has never licensed or permitted Shazam to practice any of the legal rights granted under the '552 patent.

18. On information and belief, Shazam infringes Claim 9 of the '552 patent, under 35 U.S.C. §§ 271(a) (direct infringement), (b) (inducing infringement), and/or (c) (contributory infringement).

19. All conditions precedent to this cause of action have occurred or been performed.

COUNT TWO – PATENT INFRINGEMENT

United States Patent No. 7,587,602

20. Digimarc incorporates the allegations stated in Paragraphs 1-14 as if fully set forth herein.

21. Digimarc is the sole owner of the entire right, title, and interest in United States Patent No. 7,587,602. A true copy of the '602 patent is attached as Exhibit B.

22. Digimarc has never licensed or permitted Shazam to practice any of the legal rights granted under the '602 patent.

23. On information and belief, Shazam infringes Claims 5, 10, 13, and 17 of the '602 patent, under 35 U.S.C. §§ 271(a) (direct infringement), (b) (inducing infringement), and/or (c) (contributory infringement).

24. All conditions precedent to this cause of action have occurred or been performed.

COUNT THREE – PATENT INFRINGEMENT

United States Patent No. 7,590,259

25. Digimarc incorporates the allegations stated in Paragraphs 1-14 as if fully set forth herein.

26. Digimarc is the sole owner of the entire right, title, and interest in United States Patent No. 7,590,259. A true copy of the '259 patent is attached as Exhibit C.

27. Digimarc has never licensed or permitted Shazam to practice any of the legal rights granted under the '259 patent.

28. On information and belief, Shazam infringes Claims 32, 48, and 61 of the '259 patent, under 35 U.S.C. §§ 271(a) (direct infringement), (b) (inducing infringement), and/or (c) (contributory infringement).

29. All conditions precedent to this cause of action have occurred or been performed.

PRAYER FOR RELIEF

WHEREFORE, Digimarc prays for the following relief:

30. A judicial determination that Shazam has infringed the above-identified claims of the '552, '602, and '259 patents;

31. An award of damages and ongoing royalty for Digimarc;

32. A permanent injunction against defendant Shazam from committing further acts of infringement; and

33. Such other relief as this Court deems just and equitable.

DEMAND FOR JURY TRIAL

Pursuant to Fed. R. Civ. P. 38 and Local Rule 38.1, Plaintiff Digimarc demands a jury trial on all issued raised in this Complaint triable to a jury.

Respectfully submitted,

Dated: November 16, 2009 By: _____

Stephen J. Joncus, OSB #013072
Email: stephen.joncus@klarquist.com
John D. Vandenberg, OSB #893755
Email: john.vandenberg@klarquist.com
KLARQUIST SPARKMAN, LLP
121 S.W. Salmon Street, Suite 1600
Portland, Oregon 97204
Telephone: 503-595-5300
Facsimile: 503-595-5301

Attorneys for Plaintiff Digimarc Corporation

US007349552B2

(12) **United States Patent**
Levy et al.

(10) **Patent No.: US 7,349,552 B2**
(45) **Date of Patent: Mar. 25, 2008**

(54) **CONNECTED AUDIO AND OTHER MEDIA OBJECTS**

(75) Inventors: **Kenneth L. Levy**, Stevenson, WA (US); **Geoffrey B. Rhoads**, West Linn, OR (US)

(73) Assignee: **Digimarc Corporation**, Beaverton, OR (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 933 days.

(21) Appl. No.: **10/338,032**

(22) Filed: **Jan. 6, 2003**

(65) **Prior Publication Data**

US 2003/0167173 A1 Sep. 4, 2003

Related U.S. Application Data

(60) Division of application No. 09/563,664, filed on May 2, 2000, now Pat. No. 6,505,160, which is a continuation-in-part of application No. 09/476,686, filed on Dec. 30, 1999, and a continuation-in-part of application No. 08/746,613, filed on Nov. 12, 1996, now Pat. No. 6,122,403, which is a continuation-in-part of application No. 08/649,419, filed on May 16, 1996, now Pat. No. 5,862,260, and a continuation-in-part of application No. PCT/US96/06618, filed on May 7, 1996, and a continuation-in-part of application No. 08/508,083, filed on Jul. 27, 1995, now Pat. No. 5,841,978.

(60) Provisional application No. 60/134,782, filed on May 19, 1999.

(51) **Int. Cl.**
G06K 9/00 (2006.01)

(52) **U.S. Cl.** **382/100**; 382/313

(58) **Field of Classification Search** 382/100, 382/232, 313, 317; 713/176; 380/210, 287, 380/54; 455/3.01, 3.06; 725/62
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

3,810,156 A 5/1974 Goldman ..... 340/347 AD
3,919,479 A 11/1975 Moon et al. ..... 179/1 SB
4,071,698 A 1/1978 Barger, Jr. et al. ..... 179/2 R

(Continued)

FOREIGN PATENT DOCUMENTS

EP 0161512 B1 7/1989

(Continued)

OTHER PUBLICATIONS

U.S. Appl. No. 60/257,822, filed Dec. 21, 2000, Aggson et al.

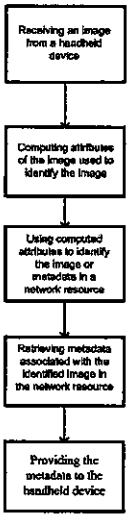
(Continued)

Primary Examiner—Andrew W. Johns

(57) **ABSTRACT**

Media objects are transformed into active, connected objects via identifiers embedded into them or their containers. In the context of a user's playback experience, a decoding process extracts the identifier from a media object and possibly additional context information and forwards it to a server. The server, in turn, maps the identifier to an action, such as returning metadata, re-directing the request to one or more other servers, requesting information from another server to identify the media object, etc. The linking process applies to broadcast objects as well as objects transmitted over networks in streaming and compressed file formats.

15 Claims, 3 Drawing Sheets



U.S. PATENT DOCUMENTS

4,230,990	A	10/1980	Lert, Jr. et al.	455/67
4,284,846	A	8/1981	Marley	179/1 SE
4,432,096	A	2/1984	Bunge	381/43
4,450,531	A	5/1984	Kenyon et al.	364/604
4,495,526	A	1/1985	Baranoff-Rossine	360/15
4,499,601	A	2/1985	Matthews	455/166
4,511,917	A	4/1985	Köhler et al.	358/84
4,547,804	A	10/1985	Greenberg	358/142
4,634,966	A	1/1987	Nakatani et al.	324/77 B
4,639,779	A	1/1987	Greenberg	358/142
4,677,466	A	6/1987	Lert, Jr. et al.	358/84
4,682,370	A	7/1987	Matthews	455/166
4,697,209	A	9/1987	Kiewit et al.	358/84
4,739,398	A	4/1988	Thomas et al.	358/84
4,776,017	A	10/1988	Fujimoto	381/43
4,805,020	A	2/1989	Greenberg	358/147
4,807,031	A	2/1989	Broughton et al.	358/142
4,843,562	A	6/1989	Kenyon et al.	364/487
4,858,000	A	8/1989	Lu	358/84
4,888,798	A	12/1989	Earnest	380/4
4,918,730	A	4/1990	Schulze	381/43
4,931,871	A	6/1990	Kramer	358/142
4,945,412	A	7/1990	Kramer	358/142
4,967,273	A	10/1990	Greenberg	358/142
4,972,471	A	11/1990	Gross et al.	380/3
5,019,899	A	5/1991	Boles et al.	358/84
5,023,907	A	6/1991	Johnson et al.	380/4
5,023,929	A	6/1991	Call	455/2
5,031,228	A	7/1991	Lu	382/38
5,134,719	A	7/1992	Mankovitz	455/154.1
5,200,822	A	4/1993	Bronfin et al.	358/142
5,210,820	A	5/1993	Kenyon	395/2
5,214,792	A	5/1993	Alwadish	455/45
5,251,301	A	10/1993	Cook	395/200
5,276,629	A	1/1994	Reynolds	364/487
5,303,393	A	4/1994	Noreen et al.	455/3.2
5,319,453	A	6/1994	Copriviza et al.	348/6
5,371,551	A	12/1994	Logan et al.	348/571
5,400,261	A	3/1995	Reynolds	364/487
5,410,598	A	4/1995	Shear	380/4
5,436,653	A	7/1995	Ellis et al.	348/2
5,437,050	A	7/1995	Lamb et al.	455/2
5,481,294	A	1/1996	Thomas et al.	348/1
5,486,686	A	1/1996	Zdybel, Jr. et al.	235/375
5,491,838	A	2/1996	Takahisa et al.	455/66
5,504,518	A	4/1996	Ellis et al.	348/2
5,539,635	A	7/1996	Larson, Jr.	364/401 R
5,564,073	A	10/1996	Takahisa	455/66
5,572,246	A	11/1996	Ellis et al.	348/2
5,572,653	A	11/1996	DeTemple et al.	395/501
5,574,519	A	11/1996	Manico et al.	396/429
5,574,962	A	11/1996	Fardeau et al.	455/2
5,577,249	A	11/1996	Califano	395/611
5,577,266	A	11/1996	Takahisa et al.	455/66
5,579,124	A	11/1996	Aijala et al.	386/96
5,579,537	A	11/1996	Takahisa	455/66
5,581,658	A	12/1996	O'Hagan et al.	395/22
5,581,800	A	12/1996	Fardeau et al.	455/2
5,584,070	A	12/1996	Harris et al.	455/346
5,612,729	A	3/1997	Ellis et al.	348/2
5,621,454	A	4/1997	Ellis et al.	348/2
5,629,980	A	5/1997	Stefik et al.	380/4
5,640,193	A	6/1997	Wellner	348/7
5,661,787	A	8/1997	Pocock	379/101.01
5,663,766	A	9/1997	Sizer, II	348/473
5,671,267	A	9/1997	August et al.	379/61
5,708,478	A	1/1998	Tognazzini	348/552
5,721,827	A	2/1998	Logan et al.	395/200.47
5,732,216	A	3/1998	Logan et al.	395/200.33
5,737,025	A	4/1998	Dougherty et al.	348/473
5,740,244	A	4/1998	Indeck et al.	380/4
5,751,854	A	5/1998	Saitoh et al.	382/218
5,761,606	A	6/1998	Wolzien	455/6.2
5,765,152	A	6/1998	Erickson	707/9
5,765,176	A	6/1998	Bloomberg	707/514
5,774,452	A	6/1998	Wolosewicz	370/212
5,774,664	A	6/1998	Hidary et al.	395/200.48
5,778,192	A	7/1998	Shuster et al.	395/200.77
5,781,914	A	7/1998	Stork et al.	707/506
5,806,031	A	9/1998	Fineberg	704/254
5,815,709	A	9/1998	Waldo et al.	395/683
5,832,119	A	11/1998	Rhoads	382/232
5,841,978	A	11/1998	Rhoads	395/200.47
5,842,162	A	11/1998	Fineberg	704/233
5,845,281	A	12/1998	Benson et al.	707/9
5,862,260	A	1/1999	Rhoads	382/232
5,875,249	A	2/1999	Mintzer et al.	380/54
5,889,868	A	3/1999	Moskowitz et al.	380/51
5,892,536	A	4/1999	Logan et al.	348/13
5,892,900	A	4/1999	Ginter et al.	395/186
5,893,095	A	4/1999	Jain et al.	707/6
5,901,224	A	5/1999	Hecht	380/4
5,902,353	A	5/1999	Reber et al.	709/219
5,903,892	A	5/1999	Hoffert et al.	707/10
5,905,248	A	5/1999	Russell et al.	235/462
5,905,800	A	5/1999	Moskowitz et al.	380/28
5,915,019	A	6/1999	Ginter et al.	380/4
5,915,027	A	6/1999	Cox et al.	380/54
5,918,223	A	6/1999	Blum et al.	707/1
5,930,369	A	7/1999	Cox et al.	380/54
5,932,863	A	8/1999	Rathus et al.	235/462.15
5,937,000	A	8/1999	Lee et al.	375/200
5,938,727	A	8/1999	Ikeda	709/218
5,940,135	A	8/1999	Petrovic et al.	348/473
5,943,422	A	8/1999	Van Wie et al.	380/9
5,945,932	A	8/1999	Smith et al.	341/51
5,974,548	A	10/1999	Adams	713/200
5,978,791	A	11/1999	Farber et al.	707/2
5,982,956	A	11/1999	Lahmi	382/306
5,983,176	A	11/1999	Hoffert et al.	704/233
5,986,651	A	11/1999	Reber et al.	345/335
5,986,692	A	11/1999	Logan et al.	348/13
5,987,509	A	11/1999	Portuesi	709/217
5,988,897	A	11/1999	Pierce et al.	400/61
5,991,500	A	11/1999	Kanota et al.	386/94
5,991,737	A	11/1999	Chen	705/26
5,991,876	A	11/1999	Johnson et al.	713/200
5,995,105	A	11/1999	Reber et al.	345/356
5,999,224	A	12/1999	Maeda et al.	348/563
6,002,443	A	12/1999	Iggulden	348/553
6,021,432	A	2/2000	Sizer, II et al.	709/217
6,028,960	A	2/2000	Graf et al.	382/203
6,058,430	A	5/2000	Kaplan	709/245
6,061,719	A	5/2000	Bendinelli et al.	709/218
6,075,568	A	6/2000	Matsuura	348/478
6,076,734	A	6/2000	Dougherty et al.	235/462.01
6,081,629	A	6/2000	Browning	382/313
6,081,827	A	6/2000	Reber et al.	709/200
6,081,830	A	6/2000	Schindler	709/204
6,084,528	A	7/2000	Beach et al.	340/825.35
6,088,455	A	7/2000	Logan et al.	380/200
6,098,106	A	8/2000	Philyaw et al.	709/238
6,121,530	A	9/2000	Sonoda	84/609
6,122,403	A	9/2000	Rhoads	382/233
6,138,151	A	10/2000	Reber et al.	709/219
6,157,721	A	12/2000	Shear et al.	380/255
6,164,534	A	12/2000	Rathus et al.	235/380
6,169,541	B1	1/2001	Smith	345/327
6,181,817	B1	1/2001	Zabih et al.	382/170
6,182,018	B1	1/2001	Tran et al.	702/66
6,185,316	B1	2/2001	Buffam	382/115
6,185,683	B1	2/2001	Ginter et al.	713/176
6,199,048	B1	3/2001	Hudetz et al.	705/23
6,199,076	B1	3/2001	Logan et al.	707/501

6,201,879	B1	3/2001	Bender et al.	382/100
6,219,787	B1	4/2001	Brewer	713/167
6,278,781	B1	8/2001	Rhoads	380/247
6,282,362	B1	8/2001	Murphy et al.	386/46
6,286,036	B1	9/2001	Rhoads	709/217
6,304,523	B1	10/2001	Jones et al.	369/30
6,311,214	B1	10/2001	Rhoads	709/217
6,314,457	B1	11/2001	Schena et al.	709/219
6,317,881	B1	11/2001	Shah-Nazaroff et al.	725/24
6,321,992	B1	11/2001	Knowles et al.	235/478.01
6,324,573	B1	11/2001	Rhoads	709/217
6,345,104	B1	2/2002	Rhoads	382/100
6,386,453	B1	5/2002	Russell et al.	235/462.01
6,389,055	B1	5/2002	August et al.	375/130
6,408,331	B1	6/2002	Rhoads	709/217
6,411,725	B1	6/2002	Rhoads	382/100
6,415,280	B1	7/2002	Farber et al.	707/2
6,433,946	B2	8/2002	Ogino	360/60
6,434,561	B1	8/2002	Durst, Jr. et al.	707/10
6,439,465	B1	8/2002	Bloomberg	235/494
6,466,670	B1	10/2002	Tsuria et al.	380/202
6,496,802	B1	12/2002	van Zoest et al.	705/14
6,505,160	B1	1/2003	Levy et al.	704/270
6,522,769	B1	2/2003	Rhoads et al.	382/100
6,523,175	B1	2/2003	Chan	725/15
6,526,449	B1	2/2003	Philyaw et al.	709/238
6,542,927	B2	4/2003	Rhoads	709/217
6,542,933	B1	4/2003	Durst, Jr. et al.	709/229
6,553,129	B1	4/2003	Rhoads	382/100
6,577,746	B1	6/2003	Evans et al.	382/100
6,614,914	B1	9/2003	Rhoads et al.	382/100
6,658,568	B1	12/2003	Ginter et al.	713/193
6,674,876	B1	1/2004	Hannigan et al.	382/100
6,674,993	B1	1/2004	Tarbouriech	455/2.01
6,681,028	B2	1/2004	Rodriguez et al.	382/100
6,697,948	B1	2/2004	Rabin et al.	713/200
6,748,360	B2	6/2004	Pitman et al.	704/270
6,748,533	B1	6/2004	Wu et al.	713/176
6,768,980	B1	7/2004	Meyer et al.	704/500
6,771,885	B1	8/2004	Agnihotri et al.	386/83
6,772,124	B2	8/2004	Hoffberg et al.	704/270.1
6,807,534	B1	10/2004	Erickson	705/51
6,829,368	B2	12/2004	Meyer et al.	382/100
6,834,308	B1	12/2004	Ikezoye et al.	709/231
6,850,252	B1	2/2005	Hoffberg	345/716
6,856,977	B1	2/2005	Adelsbach et al.	705/57
6,931,451	B1	8/2005	Logan et al.	709/231
6,941,275	B1	9/2005	Swierczek	705/26
6,965,682	B1	11/2005	Davis et al.	382/100
6,968,337	B2	11/2005	Wold	707/100
6,970,886	B1	11/2005	Conwell et al.	707/104.1
6,973,669	B2	12/2005	Daniels	725/112
6,987,862	B2	1/2006	Rhoads	382/100
6,990,453	B2	1/2006	Wang et al.	704/270
7,047,413	B2	5/2006	Yacobi et al.	713/176
7,050,603	B2	5/2006	Rhoads et al.	382/100
7,058,697	B2	6/2006	Rhoads	709/217
7,127,744	B2	10/2006	Levy	726/26
2001/0007130	A1	7/2001	Takaragi	713/186
2001/0011233	A1	8/2001	Narayanaswami	705/26
2001/0026618	A1	10/2001	Van Wie et al.	380/232
2001/0026629	A1	10/2001	Oki	382/100
2001/0031066	A1	10/2001	Meyer et al.	382/100
2001/0032312	A1	10/2001	Runje et al.	713/172
2001/0044824	A1	11/2001	Hunter et al.	709/203
2001/0051915	A1	12/2001	Ueno et al.	705/39
2001/0055391	A1	12/2001	Jacobs	380/241
2002/0010826	A1	1/2002	Takahashi et al.	711/100
2002/0021805	A1	2/2002	Schumann et al.	380/201
2002/0021822	A1	2/2002	Maeno	382/100
2002/0023020	A1	2/2002	Kenyon et al.	705/26
2002/0023148	A1	2/2002	Ritz et al.	709/219
2002/0023218	A1	2/2002	Lawandy et al.	713/176
2002/0028000	A1	3/2002	Conwell et al.	382/100
2002/0032698	A1	3/2002	Cox	707/501.1
2002/0032864	A1	3/2002	Rhoads et al.	713/176
2002/0037083	A1	3/2002	Weare et al.	381/58
2002/0040433	A1	4/2002	Kondo	713/180
2002/0044659	A1	4/2002	Ohta	380/241
2002/0048224	A1	4/2002	Dygert et al.	369/1
2002/0052885	A1	5/2002	Levy	707/200
2002/0059208	A1	5/2002	Abe et al.	707/3
2002/0068987	A1	6/2002	Hars	700/94
2002/0069107	A1	6/2002	Werner	705/14
2002/0072982	A1	6/2002	Barton et al.	705/26
2002/0072989	A1	6/2002	Van de Sluis	705/26
2002/0075298	A1	6/2002	Schena et al.	345/738
2002/0082731	A1	6/2002	Pitman et al.	700/94
2002/0083123	A1	6/2002	Freedman et al.	709/203
2002/0087885	A1	7/2002	Peled et al.	713/201
2002/0088336	A1	7/2002	Stahl	84/609
2002/0099555	A1	7/2002	Pitman et al.	704/500
2002/0102966	A1	8/2002	Lev et al.	455/412
2002/0118864	A1	8/2002	Kondo et al.	382/117
2002/0126872	A1	9/2002	Brunk et al.	382/100
2002/0133499	A1	9/2002	Ward et al.	707/102
2002/0138744	A1	9/2002	Schleicher et al.	713/187
2002/0150165	A1	10/2002	Huizer	375/240.25
2002/0152388	A1	10/2002	Linnartz et al.	713/176
2002/0153661	A1	10/2002	Brooks et al.	273/288
2002/0161741	A1	10/2002	Wang et al.	707/1
2002/0168082	A1	11/2002	Razdan	382/100
2002/0174431	A1	11/2002	Bowman et al.	725/47
2002/0178410	A1	11/2002	Haitsma et al.	714/709
2002/0184505	A1	12/2002	Mihcak et al.	713/180
2003/0018709	A1	1/2003	Schrempp et al.	709/203
2003/0028796	A1	2/2003	Roberts et al.	713/193
2003/0037010	A1	2/2003	Schmelzer	705/67
2003/0051252	A1	3/2003	Miyaoku et al.	725/109
2003/0101162	A1	5/2003	Thompson et al.	707/1
2003/0120679	A1	6/2003	Kriechbaum et al.	707/102
2003/0135623	A1	7/2003	Schrempp et al.	709/227
2003/0167173	A1	9/2003	Levy et al.	704/273
2003/0174861	A1	9/2003	Levy et al.	382/100
2003/0197054	A1	10/2003	Funson	235/375
2004/0049540	A1	3/2004	Wood	709/203
2004/0145661	A1	7/2004	Murakami et al.	348/222.1
2004/0169892	A1	9/2004	Yoda	358/3.28
2004/0201676	A1	10/2004	Needham	348/207.1
2004/0223626	A1	11/2004	Honsinger et al.	382/100
2005/0043018	A1	2/2005	Kawamot	455/414.3
2005/0058319	A1	3/2005	Rhoads et al.	382/100
2005/0108242	A1	5/2005	Kalker et al.	707/10
2005/0144455	A1	6/2005	Haitsma	713/176
2005/0229107	A1	10/2005	Hull et al.	715/764
2005/0267817	A1	12/2005	Barton et al.	705/26

FOREIGN PATENT DOCUMENTS

EP	0493091	A1	7/1992
EP	0 581 317	A2	2/1994
EP	0 649 074	A1	4/1995
EP	0464328	B1	9/1995
EP	0 901 282	A2	3/1999
EP	0 953938	A2	11/1999
EP	0967803	A2	12/1999
EP	1173001	A2	1/2002
EP	1199878	A2	4/2002
JP	11265396	A	9/1999
JP	11-272287	A	10/1999
WO	WO94/00842	A1	1/1994
WO	WO97/33273	A1	9/1997
WO	WO97/41683	A1	11/1997
WO	WO 97/43736	A1	11/1997
WO	WO98/03923	A1	1/1998
WO	WO98/20675	A1	5/1998

WO	WO98/36372	A1	8/1998
WO	WO98/43237	A1	10/1998
WO	WO 99/18723	A1	4/1999
WO	WO99/35809	A1	7/1999
WO	WO 99/41900	A1	8/1999
WO	WO99/48099	A1	9/1999
WO	WO 99/57623	A2	11/1999
WO	WO00/58940	A2	10/2000
WO	WO00/70585	A1	11/2000
WO	WO00/79709	A1	12/2000
WO	WO 01/01331	A1	1/2001
WO	WO01/15021	A2	3/2001
WO	WO01/20483	A2	3/2001
WO	WO01/20609	A2	3/2001
WO	WO01/35676	A1	5/2001
WO	WO01/62004	A2	8/2001
WO	WO01/71517	A1	9/2001
WO	WO01/72030	A2	9/2001
WO	WO 0172030		9/2001
WO	WO01/75629	A1	10/2001
WO	WO0175794	A2	10/2001
WO	WO02/11123	A2	2/2002
WO	WO02/19589	A1	3/2002
WO	WO02/27600	A2	4/2002
WO	WO002082271	A1	10/2002

OTHER PUBLICATIONS

U.S. Appl. No. 60/263,490, filed Jan. 22, 2001, Brunk et al.
U.S. Appl. No. 09/511,632, filed Feb. 11, 2000, Ikezoye et al.
U.S. Appl. No. 60/232,618, filed Sep. 14, 2000, Cox.
U.S. Appl. No. 60/175,159, filed Jan. 7, 2000, Derose et al.
U.S. Appl. No. 60/178,028, filed Jan. 26, 2000, Meyer et al.
Depovere, et al., "The VIVA Project: Digital Watermarking for Broadcast Monitoring," 1999 IEEE. pp. 202-205.
Kim, W.G et al., "A watermarking Scheme for Both Spatial and Frequency Domain to Extract te Seal Image Without the Original Image," Proc. 5.sup.th Int. Symp. on Signal Processing and its Applications, Aug. 1999, pp. 293-296.
Lu et al., "Highly Robust Image Watermarking Using Complementary Modulations," Proc. 2.sup.nd Information Security Workshop, LNCS vol. 1729, Nov. 1999, pp. 136-153.
Mintzer et al., "If One Watermark is Good, Are More Better?," Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Mar. 1999. pp. 2067-2069.
Ohbuchi et al., "A Shape-Preserving Data Embedding Algorithm for NURBS Curves and Surfaces," Proc. Computer Graphics International (CGI '99), Jun. 1999, pp. 180-187.
Onishi et al., "A Method of Watermaking with Multiresolution Analysis and Pseudo Noise Sequences," Systems and Computers in Japan, vol. 29, No. 5, May 1998, pp. 11-19.
Voyatzis et al., "The Use of Watermarks in the Protection of Digital Multimedia Products," Proc. of the IEEE, vol. 87, No. 7, Jul. 1999. pp. 1197-1207.
Yi et al., "Agent-Based Copyright Protection Architecture for Online Electronic Publishing," Proc. SPIE vol. 3657: Security and Watermarking of Multimedia Contents, Jan. 1999, pp. 484-493.
Zhao, "A WWW Service to Embed and Prove Digital Copyright Watermarks," Proc. of the Euro. Conf. on Multimedia Applications, Services and Techniques, May, 1996, 15 pages.
U.S. Appl. No. 60/134,782, filed May 19, 1999, Rhoads.
U.S. Appl. No. 09/343,104, filed Jun 29, 1999, Rodriguez et al.
U.S. Appl. No. 09/476,686, filed Dec. 30, 1999, Rhoads et al.
U.S. Appl. No. 09/531,076, filed Mar. 18, 2000, Rhoads et al.
U.S. Appl. No. 09/547,664, filed Apr. 12, 2000, Rhoads et al.
U.S. Appl. No. 09/574,726, filed May 18, 2000, Rhoads et al.
U.S. Appl. No. 09/420,945, filed Oct. 19, 1999, Kenyon et al.
U.S. Appl. No. 60/218,824, filed Jul. 18, 2000, Kenyon et al.
U.S. Appl. No. 60/155,064, filed Sep. 21, 1999, Kenyon.
U.S. Appl. No. 09/343,104, filed Jun. 29, 1999, Rodriguez et al.
U.S. Appl. No. 60/191,778, filed Mar. 24, 2000, Ramos et al.
Lin, et al., "Generating Robust Digital Signature for Image/Video Authentication," Proc. Multimedia and Security workshop at ACM Multimedia'98, Sep. 1, 1998, pp. 49-54.
Aust, D., "Augmenting Paper Documents with Digital Information in a Mobile Environment," MS Thesis, University of Dortmund, Department of Computer Graphics, Sep. 3, 1996.
Arai et al, "Retrieving Electronic Documents with Real-World Objects on InteractiveDESK," UIST '95, Nov. 14, 1995.
Arai, InteractiveDESK: A Computer-Augmented Desk Which Responds to Operations on Real Objects, CHI 95, May 7, 1995.
Worring, "Hyperdocument Generation Using OCR and Icon Detection," Proc. 3d Int. Conf. on Doc. Analysis and Recognition, Aug. 14, 1995.
Ghias et al, Query by Humming: Musical Information Retrieval In An Audio Database. In ACM Multimedia, pp. 231-236, Nov. 1995.
Kageyama et al, Melody Retrieval with Humming, Proceedings of Int. Computer Music Conference (ICMC), 1993.
Muscle Fish press release, Muscle Fish's Audio Search Technology to be Encapsulated into Informix Datablade Module, Jul. 10, 1996.
Wold et al, Content-Based Classification, Search, and Retrieval of Audio, IEEE Multimedia Magazine, Fall, 1996.
Feb. 22, 2007 Amendment and Sep. 22, 2006 Office Action, each from assignee's U.S. Appl. No. 10/338,031 (published as US 2[illegible]-0174861 A1).

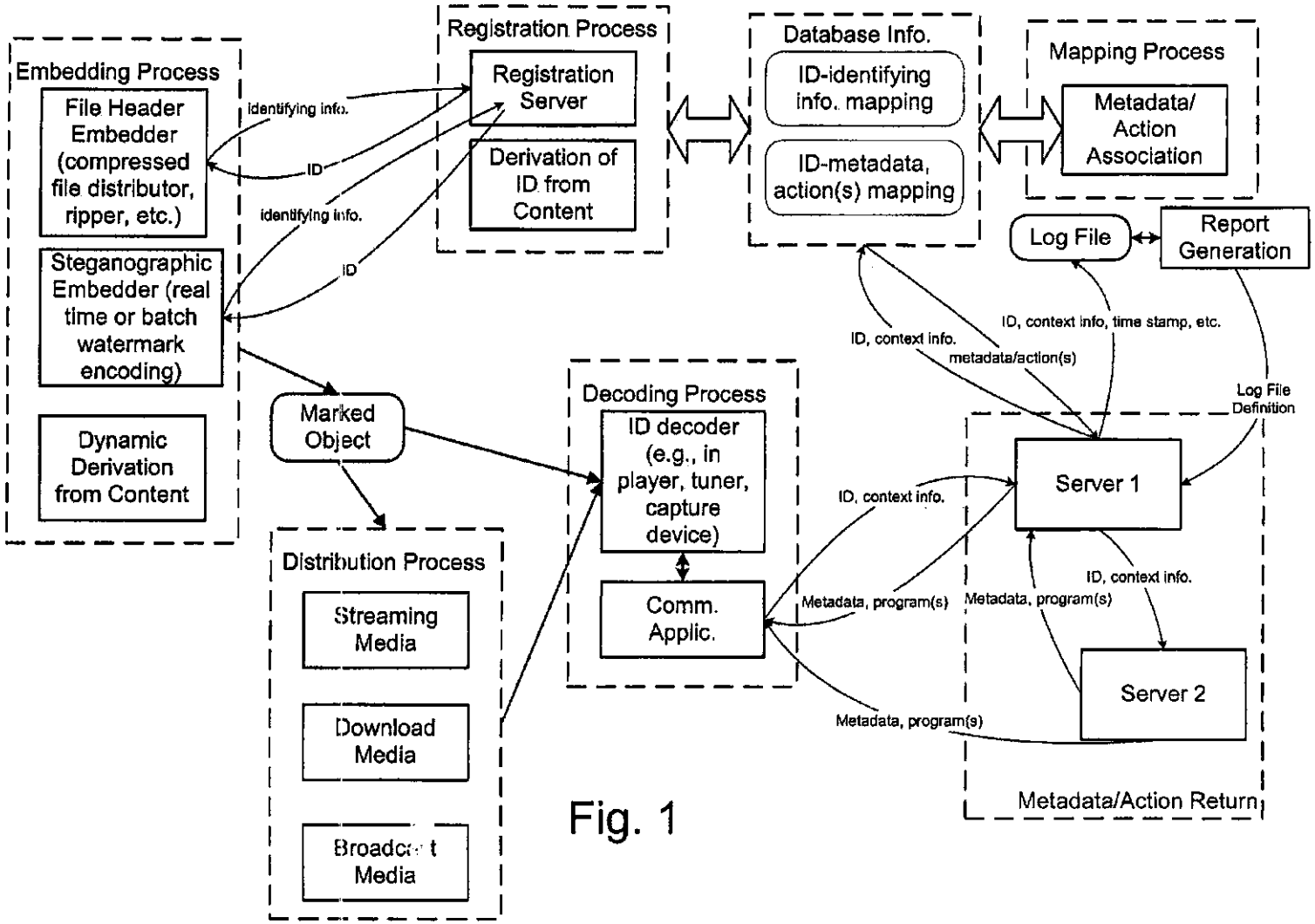

Embedding Process
File Header Embedder (compressed file distributor, ripper, etc.)
Steganographic Embedder (real time or batch watermark encoding)
Dynamic Derivation from Content
Registration Process
Registration Server
Derivation of ID from Content
Database Info.
ID-identifying info. mapping
ID-metadata, action(s) mapping
Mapping Process
Metadata/ Action Association
Log File
Report Generation
Identifying info.
ID
Identifying info.
ID
ID, context info.
metadata/action(s)
ID, context info, time stamp, etc.
Log File Definition
Marked Object
Decoding Process
ID decoder (e.g., in player, tuner, capture device)
Comm. Applic.
ID, context info.
Server 1
Metadata, program(s)
Metadata, program(s)
ID, context info.
Server 2
Metadata, program(s)
Metadata/Action Return
Distribution Process
Streaming Media
Download Media
Broadcast Media

Fig. 1

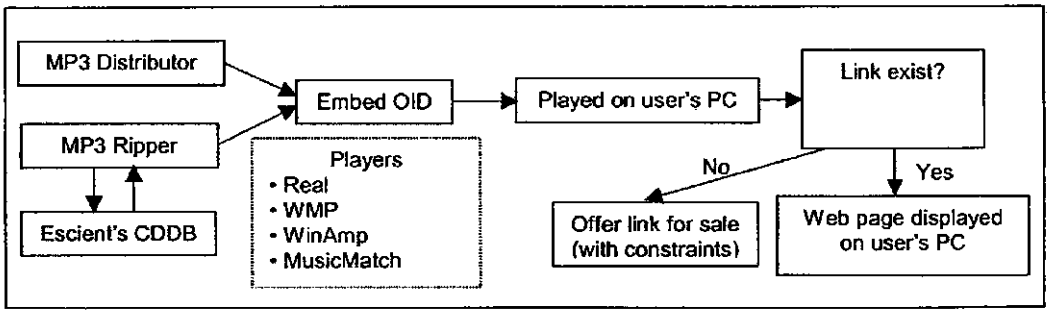

MP3 Distributor
MP3 Ripper
Escient's CDDB
Embed OID
Players
• Real
• WMP
• WinAmp
• MusicMatch
Played on user's PC
Link exist?
No
Yes
Offer link for sale (with constraints)
Web page displayed on user's PC

Fig. 2

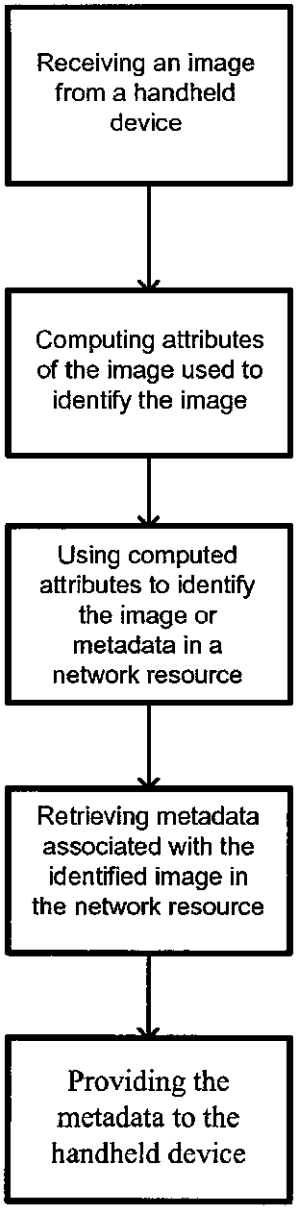

Receiving an image from a handheld device
Computing attributes of the image used to identify the image
Using computed attributes to identify the image or metadata in a network resource
Retrieving metadata associated with the identified image in the network resource
Providing the metadata to the handheld device

Fig. 3A

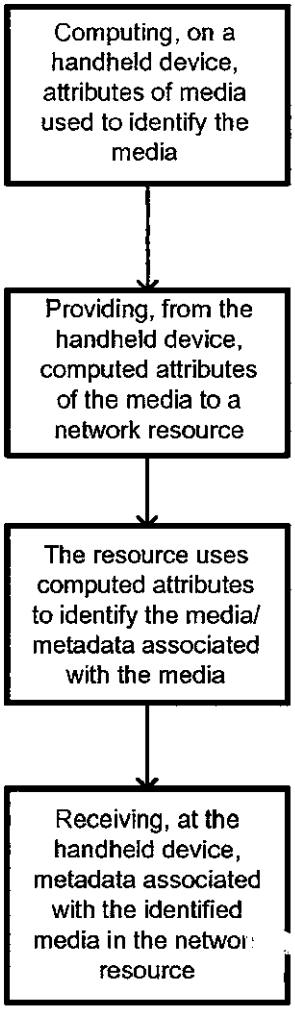

Computing, on a handheld device, attributes of media used to identify the media
Providing, from the handheld device, computed attributes of the media to a network resource
The resource uses computed attributes to identify the media/ metadata associated with the media
Receiving, at the handheld device, metadata associated with the identified media in the networ resource

Fig. 3B

CONNECTED AUDIO AND OTHER MEDIA OBJECTS

RELATED APPLICATION DATA

This patent application is a divisional of U.S. patent application No. 09/563,664, filed May 2, 2000 (now U.S. Pat. No. 6,505,160), which is a continuation in part of U.S. patent application No. 09/476,686, filed Dec. 30, 1999, which claims priority to U.S. Provisional Application No. 60/134,782, filed May 19, 1999.

U.S. patent application Ser. No. 09/563,664 is also a continuation in part of U.S. patent application Ser. No. 08/746,613 filed Nov. 12, 1996 (now U.S. Pat. No. 6,122,403), which is a continuation in part of U.S. patent application Ser. No. 08/649,419, filed May 16, 1996 (now U.S. Pat. No. 5,862,260), PCT Application PCT/US96/06618, filed May 7, 1996, and U.S. patent application Ser. No. 08/508,083, filed Jul. 27, 1995 (now U.S. Pat. No. 5,841,978).

The subject matter of the present application is related to that disclosed in U.S. Pat. No. 5,862,260, and in co-pending application Nos. 08/746,613, filed Nov. 12, 1996 (allowed); 09/343,104, filed Jun. 29, 1999; 60/164,619, filed Nov. 10, 1999; 09/476,686, filed Dec. 30, 1999; 09/503,881, filed Feb. 14, 2000; 09/525,865, filed Mar. 15, 2000; 60/191,778 filed March 24; and 09/547,664, filed Apr. 12, 2000, which are hereby incorporated by reference.

TECHNICAL FIELD

The invention related to liking audio and other multimedia data objects with metadata and actions via a communication network, e.g., computer, broadcast, wireless, etc.

BACKGROUND AND SUMMARY

Advances in computer and wireless networking, multimedia coding, and higher bandwidth communication links are creating many new ways to distribute and enjoy multimedia content, such as music and movies. Coding formats for audio like MPEG 1 Layer 3 (MP3) have already caused significant changes in music delivery to consumers. Despite the advances in technology, conte[illegible]tributors and broadcasters still need to address how to effectively promote and sell content.

This disclosure describes systems and processes for linking audio and other multimedia data objects with metadata and actions via a communication network, e.g., computer, broadcast, wireless, etc. Media objects are transformed into active, connected objects via identifiers embedded into them or their containers. These identifiers can be embedded by the owner or distributor of the media object, or automatically created from the media object. In the context of a user's playback experience, a decoding process extracts the identifier from a media object and possibly additional context information and forwards it to a server. The server, in turn, maps the identifier to an action, such as returning metadata, re-directing the request to one or more other servers, requesting information from another server to identify the media object, etc. If the identifier has no defined action, the server can respond with an option for the user to buy the link and control the resulting action for all objects with the current identifier. The linking process applies to broadcast objects as well as objects transmitted over networks in streaming and compressed file formats.

Further features will become apparent with reference to the following detailed description and accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a diagram illustrating examples of media object linking processes and systems.

FIG. 2 is a diagram illustrating media object linking applications.

FIGS. 3A and 3B are diagrams illustrating implementations of invention as currently claimed.

DETAILED DESCRIPTION

Linking Audio and Other Media Objects via Identifiers

The following sections describe systems and processes for linking audio and other media objects to metadata and actions via an identifier. For the sake of illustration, the disclosure focuses on a specific media type, namely audio signals (e.g., music, sound tracks of audio visual works, voice recordings, etc.). However, these systems, their components and processes apply to other types of media signals as well, including video, still images, graphical models, etc. As described further below, an identifier attached to an audio signal is used to connect that signal with metadata and/or programmatic or device actions. In the context of this document, the terms "media object" and "audio object" refer to an electronic form of a media signal and audio signal, respectively. The linking of media signals applies to objects that are transmitted over wire networks (such as a computer network), wireless networks (such as a wireless telephone network), and broadcast (AM, FM, digital broadcast, etc.).

There are a number of ways to associate an identifier with an audio object. One way to associate the identifier is to insert it in the form of a numeric or alphanumeric code (e.g., binary or M-ary code) in the electronic file in which the audio is stored. Another way to associate the identifier is to embed it as auxiliary data in the audio signal using steganographic methods, such as digital watermarking or other data hiding techniques. Yet another way is to derive the identifier from the audio signal, the table of contents, the file system [illegible]cture, or its container (e.g., an electronic file or physical package for data like flash memory, Digital Versatile Disk (DVD), minidisk, or compact disk (CD). The physical media may have identifying characteristics, such as a unique identifier or encoded metadata, or other attributes from which an identifier can be derived (e.g., CD disk wobble).

When the identifier is associated with metadata or actions, it transforms the media object into a "linked" object. The identifier travels with the object through distribution, including in some cases, through physical distribution in packaged media and through electronic distribution (broadcast or network communication). The identifier may travel within the same band as the audio object, such as a watermark, or via a separate band, such as a file header or footer or separate broadcast band. A decoding device or programmatic process extracts the identifier from the object and uses it to retrieve related data or actions ("metadata") In the case of an audio object, like a song, the metadata typically includes the title, artist, lyrics, copyright owner, sound recording owner, information about buying or sampling opportunities and URLs to this type of data as well as web sites and other programs and devices. Linked actions include device or programmatic processes for electronically establishing a license, transfer-

ring content (either streaming or download), sending an email, recording marketing data about a transaction, etc. The identifier allows a fan of a particular type of music or artist to get more information about the music and to buy more music. From the perspective of the artists and record labels, the identifier provides an additional opportunity to promote their music and sell content, concert tickets, etc.

In addition, in some implementations where identifier linking transactions are monitored, it enables the vendors of music to gather data about electronic transactions triggered by the link. For example, users of information may choose to provide information about themselves when they register their decoding device or software with the system. A user ID or other context information may then be recorded when the identifier is extracted and used to trigger a transaction. Many entities involved in the distribution of media signals can benefit from the linking capability. Artists can link their music to information about themselves and provide electronic buying opportunities for music, concert tickets, clothing, etc. Rights holding organizations can use the link to inform users about itself and licensing opportunities. In some cases, the link may also be used to monitor playing and distribution of copies of the music. Record labels can link their music to information about the artist, the label, electronic buying opportunities, etc. Electronic retailers can increase sales by linking users to opportunities to sample and buy additional music (via download or streaming delivery over a wire or wireless network). Conventional brick and mortar retailers can use linking to provide information about the music and to provide buying opportunities. Radio stations and other broadcasters can use the linking capability to bring users to their web sites, creating advertising revenue, to provide electronic buying opportunities for music, concert tickets, clothing items, etc. These and other forms of linked metadata and actions may be implemented in various combinations in different application scenarios.

Depending on the application, the identifier may identify the media object in which it is embedded, or entities, things or actions other than that particular media object. One type of identifier is an object ID that identifies an audio object. This identifier may be a number associated with the object, such as its International Standard Recording Code (ISRC). Another type of identifier is distributor ID that identifies the distributor of the audio object. Another type of identifier is a broadcaster ID that identifiers the broadcaster of the audio object. Of course, more than one identifier may be encoded into an audio object or its container. In the event that an object ID is not encoded with an audio object, but instead, a distributor or broadcaster identifier is encoded with the object, other context information, such as the time of play back or distribution, location of distribution, etc. may be used to identify the audio object as part of the linking process. An example is a radio station that marks its broadcasts with a station ID and maintains a playlist database with the air times of each audio object. At decoding time, the station ID is extracted and used along with context information such as the air time of the audio object to look up the audio object or its corresponding metadata and actions. This approach enables the linking system to provide audio object specific metadata or actions even without requiring a unique object identifier in every audio object.

System Implementation

FIG. **1** is a diagram of a system configuration of linked media objects. In this configuration, an identifier links audio objects to metadata via an electronic network, such as the Internet, a wireless network, or a broadcast network. As depicted in FIG. **1**, an embedding process may be used to encode an identifier in an audio object or its container. In some cases, an embedding process encodes the identifier in the audio file (e.g., a tag in a file header or footer), in the audio signal (a digital watermark), or in the physical packaging. The identifier may also be derived as a function of the audio signal or other information in the file or physical packaging (e.g., track information on a CD). In the case of dynamically derived identifiers, an embedding process is not necessary because the identifier can be derived from the content at decoding time.

In some application scenarios, the embedding process interacts with a registration process to get an identifier. The embedding process provides information about the object (e.g., a title and artist name, an ISRC, name of distributor, etc.). In response, the registration process provides an identifier and stores a database record of the association between identifier and the object or other information used in decoding to identify the object, such as its distributor or broadcaster. The registration process may be used to assign an identifier to an audio object and to distributors or broadcasters of audio objects. The embedding and registration processes may occur before the audio object is distributed to consumers, or sometime thereafter, such as when a user transfers (e.g., "rips" an a media object from one format to another (e.g., a packaged format to an electronic file format such as a compressed file format).

Once registered, an interactive or automated mapping process associates the identifier with data or actions. The registration process creates a database of identifiers and associates the identifiers with corresponding media objects, distributors, broadcasters, etc. The mapping process associates the identifiers with corresponding metadata or actions.

Once associated with an audio object and metadata, the identifier transforms the audio object into a linked object. The identifier remains with the object through distribution, although some embedding processes are more robust than others to intentional or unintentional distortion/removal of the identifier. There a variety of different distribution scenarios. Some examples depicted in FIG. 1 include transferring an audio object over a computer network, streaming the object over a computer network, or broadcasting it (e.g., AM/FM broadcasting, digital broadcasting, broadcasting over wireless carriers, etc.). Whatever the distributions process, a user ultimately receives the linked object in a player, tuner, or capture device.

To activate the linked object, a decoding process extracts the identifier and uses it to access associated data or actions. The decoding process may be implemented as a separate program or device, or integrated into a player, tuner, or some other capture device, such as listening devices that convert ambient audio waves to an electronic signal and then extract the identifier from the signal.

In the configuration shown in FIG. **1**, the decoding process forwards the extracted identifier to a communication application, which in turn, forwards it in a message to a server. The decoding process or the communication application may add additional context information to the message sent to the server. The context information may relate to the user, the user's device, the attributes of the session (time of playback, format of playback, type of distribution (e.g., broadcast or transmitted audio file), etc.) Based on identifier and optional context information, the server determines an associated action to perform, such as re-directing an identifier or context data to another server, returning metadata (including programs, content, etc.), downloading content, logging a transaction record. To find the associated

action or actions, the server maps the identifier to actions based on the information established in the mapping process. The server may: 1) look up the data and actions in a local database stored in its memory subsystem; 2) route the identifier to one or more other servers via the network, which in turn look up related actions and data associated with the identifier; or 3) perform some combination of actions 1 and 2.

In the first case, server **1** returns data or actions associated with the identifier. The server may look up related data based on the identifier alone, or based on the identifier and other context information. Context information may be information provided by the user, by the user's computer or device, or by some other process or device. In the second case, the server looks up one or more addresses associated with the identifier and forwards the identifier and/or possibly other context data to secondary servers at these addresses via conventional networking protocols. Again, this context data may include data from the user, the user's computer, some other device or database. For example, server **1** might query a remote database for instructions about how to process an identifier. These instructions may specify data to return to the communication application or to forward to another server, which in turn, looks up associated data and returns it to the communication application. A server may return data that an audio player displays to the user or uses to control rendering of the content. For example, the server can tell the player that the object contains inappropriate content for children. The player or user can make decisions about whether or how to play the material based on this information.

Both the server and the player can adopt a set of rules. The server rules may be used to control what the server returns in response to an identifier and context data. The player rules may be used to control what the player displays to the user or how it renders the content based on data returned from a server.

Either the first server, or a server one or more levels of indirection from the identifier may return data and programmatic actions to a player via the communication application. Each server in these levels of indirection receives a database key, such as an identifier or context information, from the previous server, and uses it to look up corresponding actions. These actions may include returning data or programs to the communication application or to previous servers in the routing path of the message from the communication application. Also, the servers may route requests for information or actions to other servers. The server or servers may return data or perform actions in response to the identifier (or other context data) that do not directly impact the decoding process, or the device in which it operates.

The system depicted in FIG. **1** allows several different interested parties to establish services linked via the identifier. For example, server **1** can be configured to provide generic promotional and/or licensing information associated with an identifier. If the content owner, distributor, retailer, artist or other related party wishes to provide information or services for a connected object, then server **1** may also route the identifier for that object, and possibly context information, the address of the communication application, and instructions, to servers maintained by these entities. These servers, in turn, provide promotional, sales, or licensing information, and electronic buying or licensing opportunities specific to that entity back to the consumer over the network via the communication application.

In the context of a network configuration, Internet protocols may be used to return data to the communication application or to the device or system in which it operates. The communication application may be implemented in a web browser, such as Internet Explorer or Netscape Navigator. Examples of ways of exchanging information between a client player and a server include returning a web page with metadata and program scripts designed to run on the end user's system. The metadata itself may include active links, such as URLs to other network resources, such as a web site or some other network service. The path of the identifier from the decoding process, and the return path from a server to the communication application may include one or more hops through a wire or wireless connection using standard wire and wireless communication protocols like TCP/IP, HTTP, XML, WAP, Bluetooth, etc. In addition, data returned to the user may be routed through one or more servers that may forward the data, and in some cases, augment the data or modify it in some fashion.

FIG. **2** is a diagram illustrating applications of the system depicted in FIG. **1**. In the application scenarios depicted in FIG. **2**, an embedding process encodes an object identifier (OID) into an audio file, such as an ID3 tag in the header of an MP3 file or audio frame headers in the MP3 file. FIG. **2** shows two embedding scenarios. The first is an MP3 distributor that embeds OIDs in MP3 files before transmitting them over a network, such as the Internet, typically via a web site interface. The second is a file ripping process where a programmed computer or other device extracts an audio object from packaged media such as a CD and converts it into a coded file format like MP3. In the latter case, the ripping process may extract metadata from the CD, such as the table of contents, and use this metadata as a key to a database (CDDB) to get information about the songs on the CD, such as title, artists, etc. The table of contents or other metadata from a package medium, such as optical or magnetic storage or flash memory, may be hashed into an index to a database entry that stores information about the media signal stored on the medium. The ripping process uses the information returned from the database to identify the audio objects on the packaged media so that they can be associated with an OID. This is an example of identifying information used to associate an OID with an audio object. As part of the coding process, the ripping process inserts the OID in the file header of the MP3 file.

Later, when a user opens or plays the marked MP3 in a player, such as a software player like the real player, Liquid Audio player, Windows Media Player (WMP), WinAmp, MusicMatch, etc., a plug-in software module in the player extracts the OID and forwards it to a server via an Internet connection. The plug-in may establish its own Internet connection, or pass the OID to an Internet Browser, which in turn, establishes a connection (if one is not already present) with the server. As an intermediate step, the plug-in may display a window with user options, such as "learn more about the song", "play the song", or both. The user can then choose to get more information by actuating the first or third options in the user interface window, which cause the plug-in to forward the OID to the server.

The server then returns a web page associated with the OID, or re-directs the OID to another server (e.g., one maintained by the content distributor or owner), which in turn, returns a web page of information about the object and links to related actions (e.g., a link to a licensing server, a link to a server for buying and downloading related music etc.). The licensing server may be programmed to download software players and new music offerings compatible with those players. For instance, the licensing server may provide software for decrypting, decoding, and playing electroni-

cally distributed music according to usage rules packaged with the electronically distributed music. In this application scenario, the linking of the MP3 file enables the content owner to market music and products that promote the sale of audio objects in other formats, included formats protected with encryption, watermark copy managements schemes, etc.

In the event that a media object is not linked, the decoding and server processes can be programmed to enable the user to purchase a link for the object. For example in one scenario, the player plug-in displays a graphic for a link information indicating that the link is available after determining that an OID is not in the file. If the user clicks on the graphic, the plug-in displays more information about the procedure for purchasing or renting a link. This information may be provided in conjunction with querying the server and displaying information returned from the server, or alternatively, providing pre-programmed information incorporated into the plug-in. If the user is interested in purchasing the link, he or she can then enter input (e.g., click on a button such as "Get Link") that initiates the process of registering an OID with the object and associating metadata or actions with the OID. The process of registering the OID and associating the OID with metadata or actions may be performed as described in this document. This scenario provides yet another mechanism for transforming content into connected content.

There are many possible variations to the applications scenarios illustrated in FIG. 2. During the file ripping process (or some other embedding process), the embedder may generate a unique ID from the metadata read from the packaged media on which the media object resides. One example of such an ID is the number derived from CD metadata currently used to index information in the CDDB database. This ID may then be embedded in the audio object or its file header/footer. During OID registration, the registration process may inform the embedding process that the OID (and thus, the object for which it was derived) has not been associated with metadata or actions. In this case, the user may be given an opportunity to purchase the link, either at the time of ripping, or in the future, wherever the object travels. In the latter case, the OID in the object is associated with an option to buy the link and customize the data and/or actions associated with that link. Rather than link to promotional information, the OID gives users an option to buy or rent the link and provides them with an opportunity to customize it (e.g., linking it to a custom web site). Once customized, other users that open or play the file will then be able to link to the customized information or actions.

To assert control over the type of customization that users may perform, the registration and mapping processes can place constraints on the types of metadata and actions that users can link to a media object.

In the multimedia content industry, there are typically many rights holders and entities involved in the distribution process. This may present a conflict when linking a media object to one entity. One way to address this problem is have an object link to many different entities. For example, the server could map an OID to many entities and return links to retailers, distributors, record labels and artists. Another way to address it is to encode additional information about the distributor in the OID. For example, the OID includes fields that identify the object and its distributor. If a user activates the link to purchase products, including media objects, then the distributor name is logged with the purchase and that distributor is credited with royalties associated with the transaction. The distributor field may also be used as a key to look up the appropriate action for the OID, such as re-directing the OID to the web server of the entity associated with that OID. In this approach, even if the OID directs a user to a record label's website, the distributor field can be used to credit the distributor with a royalty for the linking transaction.

The entity responsible for maintaining a web site linked via on identifier can make deals with online resources for providing data about a media object such as lyrics, song titles, radio station play lists. The website may link to this information, access it via a database manager, etc.

File Identifiers

One form of identifier is an identifier that is inserted in an audio object file, but in a distinct field from the audio signal itself. Some examples are file headers and footers. This file identifier may be assigned before or after distribution of the audio object to consumers. In addition, it may be derived from the audio signal or other information in the file. For example, an identifier generator may derive a unique or sufficiently unique identifier from a portion of a music signal. A variety of methods for generating a unique numbers based on a unique collection of numbers may be used.

The process of embedding a file identifier may be done at the time of encoding or transcoding a file. For example, the file identifier may be inserted during a ripping process, such as when a device or programmatic process converts a song from a format stored on packaged media, like a CD or DVD, to an electronic, and compressed form, such as MP3 or some other audio codec. As another example, the file identifier may be inserted when a device or programmatic process transcodes an electronic music file from one codec format to another. Yet another example is where a file is taken from a digital or analog uncompressed format, and placed in another format for distribution.

Identifiers Embedded in Audio Signal

Another way to associate an identifier with an audio signal is to embed the identifier in the audio signal using steganographic methods, such as digital watermarking or other data hiding techniques. Many of such techniques have been developed and are described in published articles and patents. Watermarking methods are described in U.S. patent application Ser. No. 09/503,881. Other examples of methods for encoding and decoding auxiliary signals into audio signals include U.S. Pat. Nos. 5,862,260, [illegible]135 and 5,945,932. For more information on steganographic applications, see the patent applications incorporated by reference.

The steganographic embedding method may be performed in a batch process. Consider a distributor of electronic music via the Internet or some other network, or a broadcaster of music such as a radio station. In each case, the distributor and broadcaster have a collection of audio objects. The embedding process may operate on this collection of objects in a batch process by retrieving an electronic version, encoding an identifier obtained from the registration process, and returning the marked version for later distribution or broadcasting. In some cases, it is desirable to do watermark embedding in an iterative process in a studio environment to encode the watermark with an intensity that achieves desired perceptibility and robustness requirements.

The steganographic embedding method may also be performed at the time of transmission of an electronic file or broadcast of the audio object. In the case of distribution via a network such as the Internet (e.g., streaming or file download), real time embedding enables the embedding process to also embed context information that is specific to

the consumer (or the consumer's computer) that has electronically ordered the object. For example, when the user requests a file in a streaming or a compressed file format via the Internet using her browser, the distributor's server can request information (perhaps voluntary) about the user to be associated with the transmitted object. Later, the decoding process or the servers that map the identifier to actions or metadata can use this information to determine the types of information to provide or responsive action to perform.

In the case of broadcasting, real time embedding enables the identifier to be steganographically embedded throughout an electronic version of the audio signal just before, or as part of the broadcasting process.

An object or distributor ID (as well as other identifiers or context information) can be embedded in the payload of a watermark that is also used for copy control. Portion of the watermark can be used to control whether the object can be played, transferred, recorded, etc., while another part can be used to carry identifiers and other metadata for linking functions described in this document. Alternatively, entirely separate watermark encoding and decoding methods may be used for copy control and linking functions.

A watermarking process may be used to encode different watermarks in the various channels of an audio signal. Message information may be embedded in one or more channels, while synchronization or orientation signals used to detect and decode the message information may be encoded in other channels. Also, different messages (e.g., different identifiers) may be encoded in different channels. At decoding time, the different identifiers can trigger different actions or link to different data.

In broadcasting applications, an identifier may be encoded along with the broadcast of the associated media signal by modulating a subcarrier of the main carrier frequency used to transmit the media signal. The subcarrier conveys auxiliary data such as the identifier, while the main carrier conveys the associated media signal. To reduce audibility of the auxiliary data (e.g., the identifier(s)) encoded in the sub-carrier, the data can be randomized by applying it to a pseudorandom or random number by some function that may be inverted in the decoding process, e.g., multiplication or exclusive OR functions. One example of sub-carrier encoding and decoding is Active HSDS 97 developed by Seiko Corporation.

Identifiers in Digital Radio Broadcasts

Some forms of digital radio broadcasts support transmission of metadata along with media signals. This metadata can also be used to carry one or more identifiers that are mapped to metadata or actions. The metadata can be encoded at the time of broadcast or prior to broadcasting. Decoding of the identifier may be performed at the digital receiver. In particular, the digital receiver receives the broadcast data, extracts the identifier, and either automatically, or at the user's direction, forwards the identifier to a server to look up the associated metadata or action.

Dynamic Identifier Extraction from Audio Content or Related Data

As noted above, another way to associate an identifier with a corresponding audio signal is to derive the identifier from the signal. This approach has the advantage that the embedding process is unnecessary. Instead, the decoding process can generate the identifier from the audio object. In this case, the decoder computes a fingerprint of the audio signal based on a specified fingerprinting algorithm. The fingerprint is a number derived from a digital audio signal that serves as a statistically unique identifier of that signal, meaning that there is a high probability that the fingerprint was derived from the audio signal in question. One component of fingerprint algorithm is a hash algorithm. The hash algorithm may be applied to a selected portion of a music file (e.g., the first 10 seconds) to create a fingerprint. It may be applied to discrete samples in this portion, or to attributes that are less sensitive to typical audio processing. Examples of less sensitive attributes include most significant bits of audio samples or a low pass filtered version of the portion. Examples of hashing algorithms include MD5, MD2, SHA, SHA1.

As an aside, fingerprinting may also be used to determine whether an audio signal has been watermarked. The fingerprinting application can evaluate a fingerprint for a received object and compare it with one for a watermarked object (or unmarked object) to determine whether the object is likely to be watermarked. Certain fingerprints can be associated with certain types of watermark methods. Using the fingerprint, a decoding device can select an appropriate watermark decoding system for the object.

While specifically discussed in the context of audio objects, the fingerprinting process applies to other types of multimedia content as well, including still images, video, graphics models, etc. For still images and video, the identifier can be derived dynamically from a compressed or uncompressed version of the image or video signal. The fingerprinting process may be tuned to generate a specific identifier based on the type of file format. For example, the process extracts the file format from the file (e.g., from a header or footer), then uses a fingerprinting process tailored for that type of file (e.g., a hash of a compressed image or video frame). The dynamic identifier computed by this process may be associated with metadata and/or actions using the processes and systems described in this document.

Registration Process

One way to implement the registration process is to build client and server application programs that communicate over a computer network using standard network communication protocols. The client may be implemented as a software program that provides identifying information about an audio object. It can obtain the information by prompting the user for the identifying information, or from extract. it from the audio object or its container. The server may be implemented as a database management program that manages identifiers and corresponding audio objects. When queried to provide an identifier for particular identifying information, the program checks whether it has already assigned an identifier to an object based on the identifying information. If so, it returns that identifier that has already been assigned. If not, it assigns a new identifier number, creates a new entry in the database for that number and its associated identifying information.

The type of identifier used to link audio objects varies with the application. As such, the registration process may vary as well. One type of identifier is a unique identifier for an audio object. Another type of identifier is one that identifies some attribute of the audio object, but does not uniquely identify it, such as a distributor or broadcaster identifier. This type of identifier requires additional context information to uniquely identify the audio object at the time of linking it to actions or metadata. For these types of identifiers, the registration process provides information identifying the attribute of the audio object, such as its distributor or broadcaster. In response, the server provides an identifier that may be embedded in several audio objects that share that attribute.

One example is a broadcaster ID, such as a radio station ID. Audio broadcast by the radio station is embedded with this radio station ID. To identify the object, context information such as the play time captured at the tuner is used along with the radio station ID extracted from the received audio signal to identify the audio object. The decoding process forwards this information to a server. Using the radio station ID and context information, the server maps the ID to an appropriate action. This may include querying a radio station's playlist database for an object identifier based on the station ID and context information. The server can then map the object identifier to an action or metadata based on the object ID returned from the playlist database. Other scenarios are possible. For example, the server could forward the station ID, context data and decoder address to a radio station server, which in turn, looks up the appropriate action or metadata (e.g., web page) and sends it to the device that decoded the station ID.

Broadcast content can also be associated with object identifiers. One way to implement the identifier assignment process is to allocate a unique set of identifiers with each broadcaster/distributor. Those broadcasters or distributors are then free to assign the identifiers to media objects as they wish. Once they complete the identifier assignment process, they may then associate the identifiers with the metadata or actions in a mapping process.

Embedding Process

The embedding process may be integrated into a software program along with the client of the registration process described in the previous section. This integration of registration and embedding functions is particularly suited to a batch embedder, where processing time required to request an identifier is less of a concern.

In real time embedding, the identifier or identifiers are preferably available for associated audio objects before embedding begins. For example, the identifiers can be maintained in a local database on the embedding computer or device and indexed by object title. Distributor and broadcast identifiers are more straightforward because they may be applied to several different audio objects.

The embedding process may also be implemented in an embedding clearinghouse system. The embedding clearinghouse is a computer or other electronic system that analyzes media objects and embeds one or more links in the media objects. The clearinghouse may be implemented in a server on a network, such as the Internet and operate on content in a "push," "pull," or some combination of push and pull models. In the push model, users and other systems send media objects to the embedding clearinghouse for analysis and embedding. The pull model, the clearinghouse has the capability to search for and gather media objects for embedding and analysis. One example of this pull model is an Internet search process called a spider that crawls the Internet, searching for media objects to analyze and embed with one or more identifying links.

The embedding clearinghouse analyzes a media object (perhaps based on out of band data like a file header or footer) and inserts an identifier. This identifier may link to a metadata and actions, such as re-direction to a web site offering products, services, and information related to the content. The embedding clearinghouse may incorporate search engine technology to execute a key word search based on information from the media object and then associate the media object with a series of related URLs returned from the Internet search. The process may be automatic, or with some user input to select which sub-set of links should be inserted.

The embedding clearinghouse may also offer an identifier embedding services for those wanting to link their media objects with metadata, actions, etc. In this application scenario, the embedding clearinghouse may be implemented as an Internet server that is accessible via a web page using conventional network communication and web protocols. To access the server, users visit a web page using an Internet browser. In exchange for a fee, which may be tendered electronically over the Internet from the user's computer to the server, the server provides an embedding service to embed an identifier into a media object uploaded from the user via the user's computer and Internet connection. The user can select the information to associate with a media object, such as generic identifying information (e.g., title, author, owner), generic licensing information, or special information or actions. The generic information is hosted by the provider of the embedding clearinghouse server, while the special purpose information and actions are accessed through re-direction. In particular, the provider of the clearinghouse server links the embedded identifier to an address or set of addresses of servers that provide the special information or actions. Then at decoding time, the decoding process sends the identifier to the provider's server, which in turn, redirects the identifier to a secondary server or servers that provide special purpose information or actions (e.g., redirect to a web page of the content owner, download related content, provide electronic licensing services, etc.).

Decoding the ID and Embedded Context Data

The implementation details of the decoding process depend on how the identifier is encoded into an audio object or its container. In the case where the identifier is encoded in a file header or footer, the decoder may be a software program or digital hardware that parses the header/footer and forwards it to the communication application. One way to implement this type of decoder is to integrate it into a media player as a plug in program. Examples of media players include Windows Media player from Microsoft, Liquid Audio player from Liquid Audio, Winamp, Real Player from Real Networks. Preferably, the plug-in gives the user visual feedback that the identifier has been detected and displays a window with options to access more information or actions available via the link. For example, the user can be presented with a user interfaces prompting the user to click for more information or buying opportunities. If the user selects these options, the plug-in forwards the user selections and identifier to the communication application, which forwards them to the server (e.g., server **1**, FIG. **1**).

In the case where the identifier is steganographically encoded in the audio object, a corresponding decoder extracts the identifier. This type of decoder may be implemented as a plug in to a software player as described in the previous paragraph. It may also be implemented in a tuner for broadcast content, or in a listening device that captures audio from the ambient environment.

In the case where the identifier is derived from the content or container metadata, the decoder captures the pertinent portion of the audio object, and generates the identifier as described above. This type of decoder can be implemented in a software or hardware player, a tuner, etc.

The decoder may collect identifiers in response to a user request while objects containing these identifiers are being played. For example, when the user is playing music, he may like a song and want to buy it or get more information. This

feature may be implemented by building an interface that has a button or voice recognition that enables the user to request information or a buy/license opportunity. Once captured, identifiers can be forwarded along with user instructions to the appropriate server.

However, one particularly useful feature is to enable the user to fetch information and make orders from music as the music is playing. The system described previously supports this feature because the decoding process can forward the identifier or identifiers, embedded context information, or additional context information (user information, play time, broadcast type, file type, player type, operating system type) to the communication application as the music is playing. The user can trigger the linking action by pressing a "fetch" button, or saying fetch to a voice activated input device that causes the decoding device to package a message and invoke the communication application (e.g., Internet browser). In turn, the communication application forwards the message to a server that parses the message and determines the associated action.

The activation of the "fetch it" feature may be made on a handheld device that communicates with a decoding device in a tuner via a wireless connection. For example, a user may press a button on a remote control device, like a key chain, which sends a wireless signal to a receiver in the tuner. The receiver invokes the decoding process. The tuner may also send metadata from the server to the remote control device for display using a similar wireless connection. Infrared or RF transceivers, for example, may be used to communicate the data back and forth.

The decoding device may also provide continuous decoding of identifiers. When the user requests a "fetch," the identifier and context information for the current song may be forwarded to the server. Also, the decoding device may automatically fetch generic information such as song title and artist so that this information is immediately available to the user.

Another possible implementation is to temporarily buffer identifiers extracted from some predetermined number of the most recent songs, titles, etc. These identifiers can be stored along with other metadata, such as a time stamp, to inform the user when they were captured. The user can then select one or more of the items to send to the server for more information or related actions.

These features may be implemented in one or more devices. While the example above discusses a remote control device and a separate tuner with a decoder, these functions may be integrated into a single device, such as a car stereo, phone handset, personal digital assistant, and a variety of other types of players or tuners.

The identifier enables dynamic linking. Dynamic linking enables the identifier encoded with a media object to remain fixed, while the metadata or actions associated with that identifier can be changed. To change the associated metadata, the mapping process edits the identifier database to associate new metadata or actions with an identifier. The mapping process can be automated to change metadata or actions associated with an identifier at periodic intervals or in response to system events. In addition, a user may change the associated metadata or actions interactively at any time. To facilitate access to the database, a web based interface can be added to the database.

Dynamically linked data returned from a server to a player environment can be displayed to the user in a variety of ways. One way is to display it in a web page or user interface window of a player. The data can be animated by scrolling it across the visual display. The data can also be displayed in the form of HTML links, which, when activated, cause the download of other data or initiate actions, such as playing streaming content from a server.

Server Types

As discussed elsewhere, the servers used to link identifiers to actions may be programmed to provide a variety of actions including:

- returning data and HTML links (e.g., in the form of an HTML document, scripts, etc.)
- downloading media signals in streaming or file format
- performing an electronic transaction (selling products like CDs, DVDs, concert tickets, etc. via computer transaction using credit cards, digital money, etc.)
- establishing a license to use a linked media object
- re-directing to another server
- performing database look up operations for related information, links, actions
- performing database look up to uniquely identify a media object based on distributor/broadcaster ID and other context information
- creating a transaction log

This is by no means in exhaustive list. Another type of server action is to initiate a process of searching a database, a collection of databases or the Internet for additional information related to a linked media object. This type of search service may be performed continuously and the results associated with the identifier. Then, in response to a request from a decoding process, the server can return a digest of the results with links to web pages for additional information.

Communication Application

The implementation details of the communication application are highly dependent on the type of communication link and protocols used to connect the decoding process to a server. Above, an Internet browser is provided as an example. A browser may be implemented in conventional PCs, handheld devices, wireless phones, stereo systems, set top boxes, etc. However, the communication application need not be based on computer network protocols. For wireless devices, where the marked content is played on wireless carrier frequencies, the communication application can employ wireless communication technology to forward identifiers and context information to servers that map this information to actions or metadata and return it via a wireless carrier frequency to user's handset.

Tracking Transactions and Report Generation

As depicted in FIG. 1 and described above, the servers for mapping identifiers to actions may be programmed to dispense a transaction log into a log file. A report generation process can then enable users to define and request queries of data from the log file based on a particular identifier, a particular type of context information (time frame, geographic location, user demographics, etc.), a particular action, etc.

Capture Devices

As noted above, the decoding process may be implemented in a variety of devices or software that process media objects. These devices and software include programmable devices such as personal computers, personal digital assistants, telephone handsets, set-top boxes, personal stereos, hi-fi components, tuners, receivers, televisions, etc. as well as hardwired devices that may be incorporated into these systems and devices.

In some contexts, it is useful to implement a recording function. This is particularly true in devices that receive a

broadcast or stream of media content and need to capture at least a portion of it to decode an identifier. Examples of these devices are radio receivers, and wireless telephone handsets. The record function may be automatic or user activated. In the latter case, the user actuates an input device to control the record process and optionally the record duration. For example, the user may hear a song that she likes and press record. The device, in turn, records at least a part of the object that is currently being received (an audio, visual or audio visual signal). The user can then decide contemporaneously or at a later time to execute the identifier decoding process on the recorded signal. The recording function can be designed to execute for a predetermined or user specified duration.

In the case of radio and television tuners/receivers, the record function can be used to capture a media signal as it is received. In the case of a telephone handset, the record function can be used for a variety of functions, such as recording part of a telephone conversation, recording speech or other ambient audio through a microphone, or recording a media signal received by the handset via a wireless communication channel. The recordings can be compressed and stored in local memory on the device. In addition, they may be annotated with metadata about the media signal, such as a time stamp to show time of capture, a location stamp to show location of capture, metadata extracted from the object (in band or out of band data), etc. The location stamp may be provided by a global positioning device. Some wireless phone systems are capable of computing location of a telephone handset via triangulation. This location data may be used to provide geographic location coordinates or the name of nearby landmark, city name, etc.

The metadata may be displayed on a display device to help the user remember the context of a particular recording. In addition, it may be provided as context information along with an identifier to a server that links the identifier and context information to metadata or actions.

Transmarking

In some applications, it may be useful to convert auxiliary information embedded in a media signal from one format to another. This converting process is referred to as transmarking. Transmarking may include converting an out of band identifier like a tag in a header/footer to a watermark or vice versa. It may also involve converting a message in one watermark format to another. The process involves a decoding operating on an input media object, and an encoding of the decoded information into the media object. It may also involve a process for removing the mark originally in the input object to avoid interference with the newly inserted mark.

There are a variety of reasons to perform transmarking. One is to make the embedded information more robust to the types of processing that the media object is likely to encounter, such as converting from one watermark used in packaged media to another watermark used in compressed, and electronically distributed media, or a watermark used in radio or wireless phone broadcast transmission applications.

This type of transmarking process may be performed at various stages of a media object's distribution path. As suggest previously, an identifier in a watermark or file header/footer may be encoded at the time of packaging the content for distribution, either in an electronic distribution format or a physical packaged medium, such as an optical disk or magnetic memory device. At some point, the media signal may be converted from one format to another. This format conversion stage is an opportunity to perform transmarking that is tailored for the new format in terms of robustness and perceptibility concerns. The new format may be a broadcast format such as digital radio broadcast, or AM or FM radio broadcast. In this case, the identifier may be transmarked into a watermark or other metadata format that is robust for broadcast applications. The new format may be a compressed file format (e.g., ripping from an optical disk to an MP3 format). In this case, the identifier may be transmarked into a file header/footer or watermark format that is robust and compatible with the compressed file format.

The transmarking process may leave an existing embedded identifier in tact and layer an additional identifier into the media object. This may include encoding a new watermark that does not interfere with an existing watermark (e.g., insert the new watermark in unmarked portions of the media object or in a non-interfering transform domain). It may also include adding additional or new identifier tags to headers or footers in the file format.

Amplifying an Embedded Identifier

Rather than converting embedded data to another format, an amplifying process may be used to renew an identifier that has become weakened or separated due to processing of the media object in which it is embedded. In this case, a decoder and encoder pair may be used to determine the current identifier and re-encode it. Of course, the encoder can also choose to embed new or additional identifiers as well.

If the previous identifier is lost, the encoder can query an identifier database established in the registration process, passing identifying information about the media object. The database uses the identifying information to find an associated identifier and returns it to the encoder for embedding in the media object.

Concluding Remarks

Having described and illustrated the principles of the technology with reference to specific implementations, it will be recognized that the technology can be implemented in many other, different, forms. To provide a comprehensive disclosure without unduly lengthening the specification, applicants incorporate by reference the patents and patent applications referenced above. These patents and patent applications provide additional implementation details. They describe ways to implement processes and components of the systems described above. Processes and components described in these applications may be used in various combinations, and in some cases, interchangeably with processes and components described above.

The particular combinations of elements and features in the above-detailed embodiments are exemplary only; the interchanging and substitution of these teachings with other teachings in this and the incorporated-by-reference patents/applications are also contemplated.

We claim:

1. A method of linking an image to a network resource, said method comprising:

receiving an image from a handheld device;

computing attributes of the image used to identify the image, wherein said computing attributes comprises deriving an identifier of the image from the image itself;

using computed attributes of the image to identify the image or metadata associated with the image in a network resource;

retrieving metadata associated with the identified image in the network resource; and

providing the metadata to the handheld device.

2. The method of claim 1, wherein the metadata comprises metadata selected from a group of metadata comprising: audio, video, an image and a URL.

3. The method of claim 1, wherein the metadata comprises at least one action.

4. The method of claim 1, wherein the metadata comprises purchasing information.

5. A method of linking an image to a network resource, said method comprising:

computing, on a handheld device, attributes of an image used to identify the image;

providing, from the handheld device, computed attributes of the image to a network resource,

wherein the network resource uses the computed attributes of the image to identify the image or metadata associated with the image;

receiving, at the handheld device, metadata associated with the identified image in the network resource.

6. The method of claim 5, wherein the metadata comprises at least one action.

7. The method of claim 5, wherein the metadata comprises authentication information.

8. The method of claim 5, wherein said computing attributes comprises deriving an identifier of the image from the image itself.

9. A method of linking media to a network resource, said method comprising:

computing, on a handheld device, attributes of media used to identify the media;

providing, from the handheld device, computed attributes of the media to a network resource,

wherein the network resource uses the computed attributes of the media to identify the media or metadata associated with the media;

receiving, at the handheld device, metadata associated with the identified media in the network resource.

10. The method of claim 9 wherein the metadata comprises metadata selected from a group of metadata comprising: audio.

11. The method of claim 9 wherein the metadata comprises at least one action.

12. The method of claim 9 wherein the metadata comprises authentication information.

13. The method of claim 9 wherein said computing, on a handheld device, attributes comprises deriving an identifier from the media itself.

14. The method of claim 13 wherein the identifier serves as a statistically unique identifier of the media.

15. The method of claim 9 wherein the media comprises at least audio, video or imagery.

* * * * *

US007587602B2

(12) **United States Patent**
Rhoads

(10) **Patent No.: US 7,587,602 B2**
(45) **Date of Patent: Sep. 8, 2009**

(54) **METHODS AND DEVICES RESPONSIVE TO AMBIENT AUDIO**

(75) Inventor: **Geoffrey B. Rhoads**, West Linn, OR (US)

(73) Assignee: **Digimarc Corporation**, Beaverton, OR (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 646 days.

(21) Appl. No.: **11/331,430**

(22) Filed: **Jan. 11, 2006**

(65) **Prior Publication Data**

US 2006/0174348 A1 Aug. 3, 2006

Related U.S. Application Data

(62) Division of application No. 09/476,686, filed on Dec. 30, 1999.

(60) Provisional application No. 60/134,782, filed on May 19, 1999.

(51) **Int. Cl.**
H04L 9/00 (2006.01)

(52) **U.S. Cl.** **713/176**; 726/26; 726/27

(58) **Field of Classification Search** 380/54–55, 380/270, 200–203; 713/176, 165, 167, 170; 705/57–59, 51; 382/100, 115; 726/26–33
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

3,950,782	A	4/1976	Carey
4,241,415	A	12/1980	Masaki
5,119,507	A	6/1992	Mankovitz
5,249,166	A	9/1993	Hamilton
5,444,779	A	8/1995	Daniele
5,473,631	A	12/1995	Moses
5,574,962	A	11/1996	Fardeau
5,598,557	A	1/1997	Doner et al.
5,613,004	A	3/1997	Cooperman et al.
5,640,193	A	6/1997	Wellner
5,740,244	A	4/1998	Indeck et al.
5,754,981	A	5/1998	Veeneman et al.
5,761,606	A	6/1998	Wolzien
5,765,152	A	6/1998	Erickson
5,774,452	A	6/1998	Wolosewicz
5,781,914	A	7/1998	Stork
5,782,692	A	7/1998	Stelovsky
5,790,172	A	8/1998	Imanaka

(Continued)

FOREIGN PATENT DOCUMENTS

EP	1223742	7/2002

(Continued)

OTHER PUBLICATIONS

Foote, "An Overview of Audio Information Retrieval," Multimedia Systems, v.7 n.1, p. 2-10, Jan. 1999.

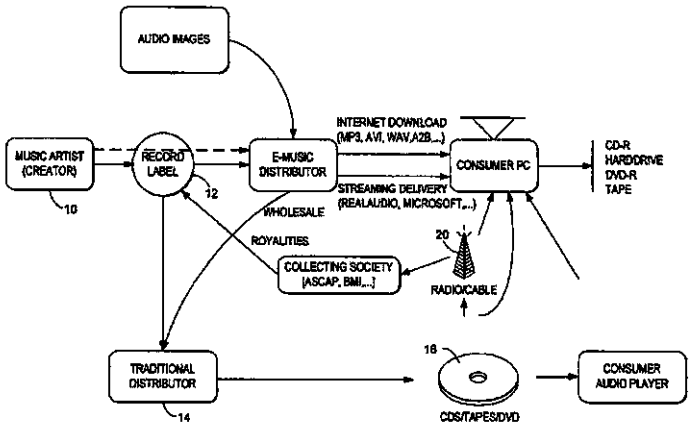
(Continued)

Primary Examiner Hosuk Song

(57) **ABSTRACT**

A portable device uses a microphone to listen to ambient audio, derives data from captured audio signals, and uses the derived data to request delivery of the audio or related information to the user's home or other location. The device is desirably pocket-sized, or suitable for carrying on a key-ring. The device may also detect a watermark signal that is present in the user's environment (e.g., played through a public address speaker system) to aid the user in recalling the context from which the audio was requested.

19 Claims, 4 Drawing Sheets



U.S. PATENT DOCUMENTS

5,801,702	A	9/1998	Dolan et al.	
5,822,432	A	10/1998	Moskowitz et al.	
5,841,978	A	11/1998	Rhoads	
5,848,413	A	12/1998	Wolff	
5,862,260	A	1/1999	Rhoads	
5,869,819	A	2/1999	Knowles et al.	
5,870,552	A	2/1999	Dozier	
5,903,892	A	5/1999	Hoffert	
5,905,248	A	5/1999	Russell et al.	
5,918,223	A	6/1999	Blum et al.	
5,923,327	A	7/1999	Smith	
5,938,727	A	8/1999	Ikeda	
5,963,957	A	10/1999	Hoffberg	
5,969,283	A	10/1999	Looney	
5,978,773	A	11/1999	Hudetz	
5,988,897	A	11/1999	Pierce et al.	
6,005,501	A	12/1999	Wolosewicz	
6,014,569	A	1/2000	Bottom	
6,064,737	A	5/2000	Rhoads	
6,076,734	A	6/2000	Dougherty et al.	
6,081,830	A	6/2000	Schindler	
6,098,106	A	8/2000	Philyaw	
6,121,530	A	9/2000	Sonoda	
6,122,526	A	9/2000	Parulski	
6,125,172	A	9/2000	August	
6,141,753	A	10/2000	Zhao et al.	
6,147,940	A	11/2000	Yankowski	
6,151,624	A	11/2000	Teare	
6,169,541	B1	1/2001	Smith	
6,185,318	B1	2/2001	Jain et al.	
6,188,010	B1	2/2001	Iwamura	
6,199,048	B1	3/2001	Hudetz et al.	
6,222,807	B1	4/2001	Min-Jae	
6,240,185	B1	5/2001	Van Wie et al.	
6,243,480	B1	6/2001	Zhao et al.	
6,249,588	B1	6/2001	Amidror	
6,278,781	B1	8/2001	Rhoads	
6,282,362	B1	8/2001	Murphy et al.	
6,286,036	B1	9/2001	Rhoads	
6,304,523	B1	10/2001	Jones	
6,338,094	B1	1/2002	Scott	
6,345,256	B1	2/2002	Milsted et al.	
6,389,055	B1 *	5/2002	August et al.	375/130
6,389,538	B1	5/2002	Gruse	
6,456,725	B1	9/2002	Cox et al.	
6,493,457	B1	12/2002	Quackenbush	
6,505,160	B1	1/2003	[illegible] et al.	
6,510,234	B1	1/2003	Cox et al.	
6,516,356	B1	2/2003	Belknap	
6,526,449	B1	2/2003	Philyaw et al.	
6,542,927	B2	4/2003	Rhoads	
6,560,349	B1	5/2003	Rhoads	
6,573,883	B1	6/2003	Bartlett	
6,628,928	B1	9/2003	Crosby	
6,640,306	B1	10/2003	Tone et al.	
6,674,993	B1	1/2004	Tarbouriech	
6,829,368	B2	12/2004	Meyer et al.	
6,931,451	B1	8/2005	Logan et al.	
6,941,275	B1	9/2005	Swierczek	
7,010,263	B1	3/2006	Patsiokas	
7,020,304	B2	3/2006	Alattar et al.	
7,043,048	B1	5/2006	Ellingson	
7,084,903	B2	8/2006	Narayanaswami	
7,164,780	B2	1/2007	Brundage et al.	
7,171,018	B2	1/2007	Rhoads et al.	
7,174,293	B2	2/2007	Kenyon	
7,224,995	B2	5/2007	Rhoads	
7,251,475	B2	7/2007	Kawamoto	
7,302,574	B2	11/2007	Bradley	
7,330,564	B2	2/2008	Brundage	
7,333,957	B2	2/2008	Levy et al.	
7,349,552	B2	3/2008	Levy	
7,370,190	B2	5/2008	Calhoon	
7,437,430	B2	10/2008	Rhoads	
7,461,136	B2	12/2008	Rhoads	
7,505,605	B2	3/2009	Rhoads	
2001/0011233	A1	8/2001	Narayanaswami	
2001/0016836	A1	8/2001	Boccon-Gibod	
2002/0033844	A1	3/2002	Levy et al.	
2002/0083124	A1	6/2002	Knox	
2002/0124024	A1	9/2002	Patterson	
2002/0166050	A1	11/2002	Takahashi	
2002/0186861	A1	12/2002	Echizen	
2003/0005135	A1	1/2003	Inoue	
2003/0011684	A1	1/2003	Narayanaswami et al.	
2003/0014496	A1	1/2003	Spencer	
2003/0167173	A1	9/2003	Levy	
2003/0174861	A1	9/2003	Levy et al.	
2003/0195033	A1	10/2003	Gazdic	
2004/0128514	A1	7/2004	Rhoads	
2004/0205055	A1	10/2004	Overton	
2004/0243634	A1	12/2004	Levy	
2005/0043018	A1	2/2005	Kawamoto	
2005/0058319	A1	3/2005	Rhoads	
2005/0065780	A1	3/2005	Wiser et al.	
2005/0091268	A1	4/2005	Meyer	
2005/0192933	A1	9/2005	Rhoads	
2005/0229107	A1	10/2005	Hull	
2005/0249374	A1	11/2005	Levy	
2006/0109984	A1	5/2006	Rhoads	
2006/0120560	A1	6/2006	Davis	
2007/0091376	A1	4/2007	Calhoon	
2007/0100757	A1	5/2007	Rhoads	
2007/0101147	A1	5/2007	Brunk	
2007/0185840	A1	8/2007	Rhoads	
2007/0189533	A1	8/2007	Rhoads	
2007/0195987	A1	8/2007	Rhoads	
2007/0276841	A1	11/2007	Rhoads	
2007/0286454	A1	12/2007	Brundage	
2008/0014917	A1	1/2008	Rhoads	
2008/0134255	A1	6/2008	Ferris	
2008/0139182	A1	6/2008	Levy	
2008/0140714	A1	6/2008	Rhoads	
2008/0260201	A1	10/2008	Rhoads	

FOREIGN PATENT DOCUMENTS

WO	WO 97/43736	11/1997
WO	WO9743736	11/1997
WO	WO0036605	6/2000
WO	WO0079709	12/2000
WO	WO 01/61508	8/2001

OTHER PUBLICATIONS

Blackburn, "A Tool for Content Based Navigation of Music," ACM Multimedia 98.

deRoure, "Multiagent System for Content Based Navigation of Music," ACM Multimedia, Oct. 1999, 4 pp.

Roy, "Wearable Audio Computer—A Survey of Interaction Techniques," MIT Media Lab, 1997.

Ghias et al, Query by Humming: Musical Information Retrieval In An Audio Database, ACM Multimedia, pp. 231-236, Nov. 1995.

* cited by examiner

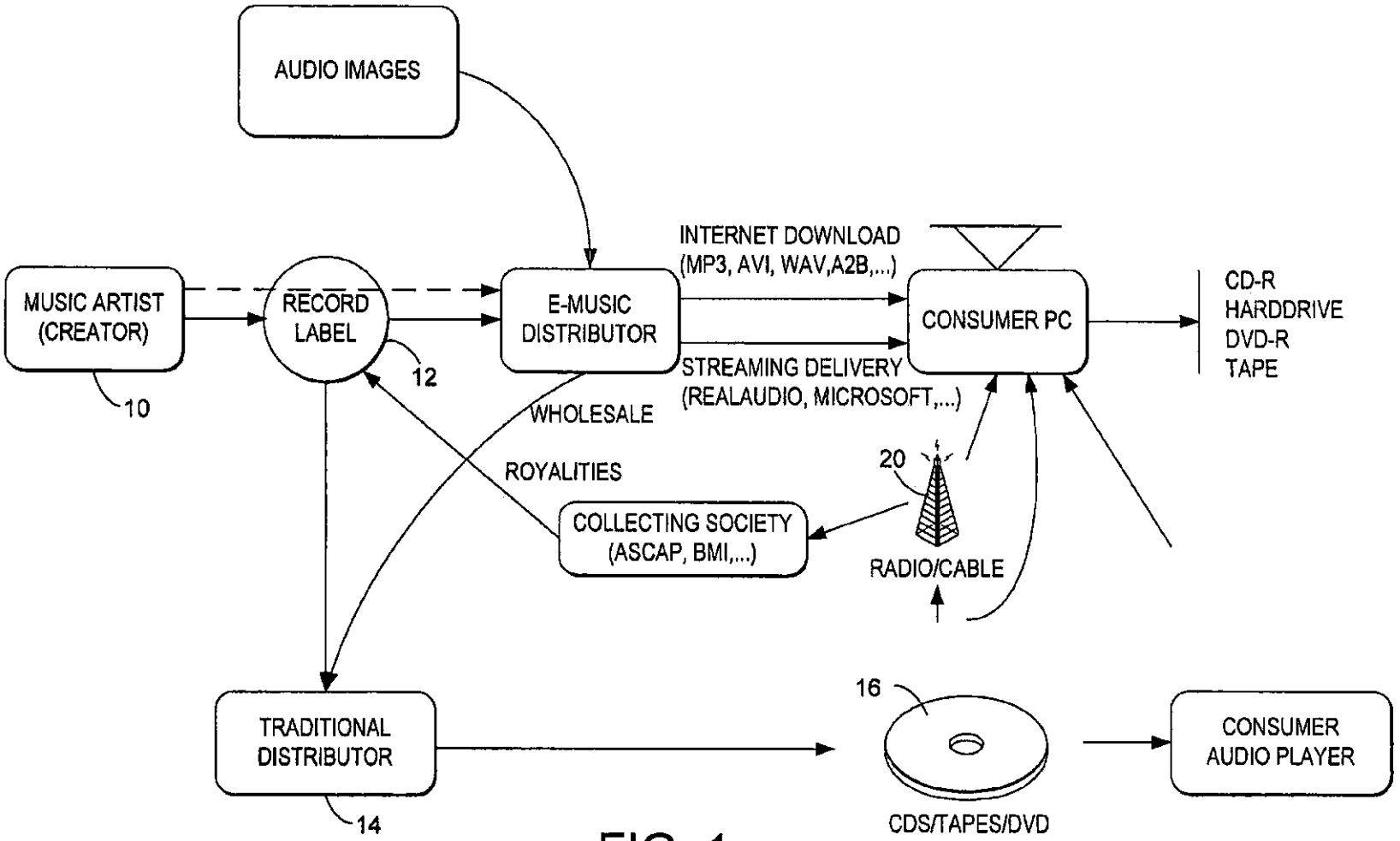

AUDIO IMAGES
MUSIC ARTIST (CREATOR)
10
RECORD LABEL
12
E-MUSIC DISTRIBUTOR
INTERNET DOWNLOAD (MP3, AVI, WAV,A2B,...)
STREAMING DELIVERY (REALAUDIO, MICROSOFT,...)
CONSUMER PC
CD-R
HARDDRIVE
DVD-R
TAPE
WHOLESALE
ROYALITIES
COLLECTING SOCIETY (ASCAP, BMI,...)
20
RADIO/CABLE
TRADITIONAL DISTRIBUTOR
14
16
CDS/TAPES/DVD
CONSUMER AUDIO PLAYER

FIG. 1

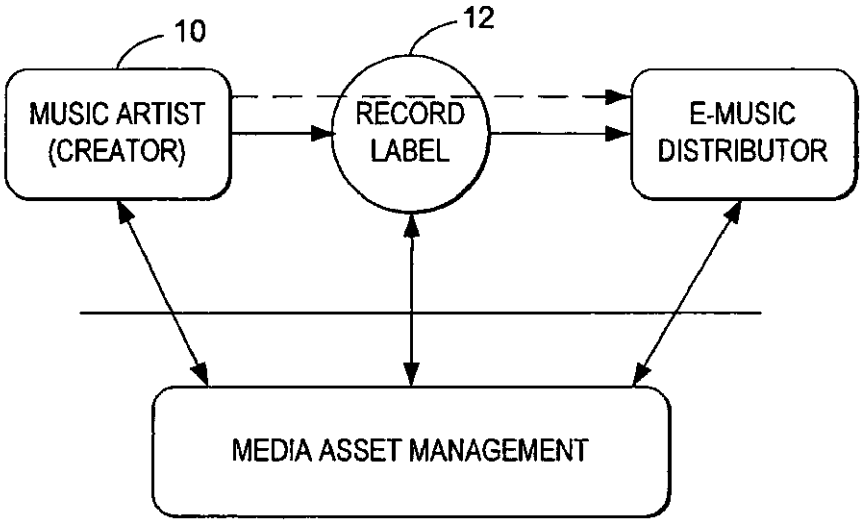

10
12
MUSIC ARTIST (CREATOR)
RECORD LABEL
E-MUSIC DISTRIBUTOR
MEDIA ASSET MANAGEMENT

FIG. 2

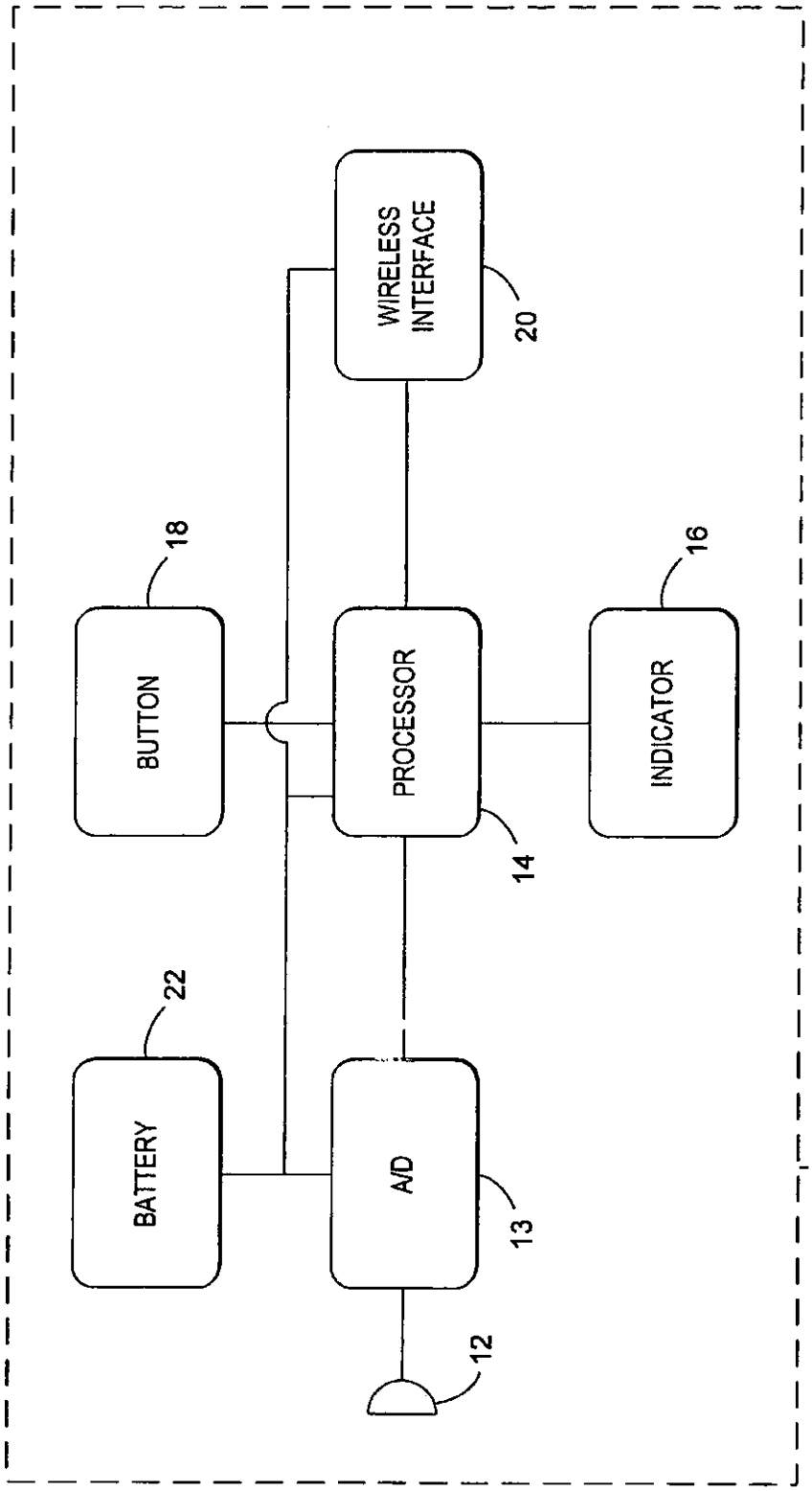

WIRELESS INTERFACE
20
18
BUTTON
PROCESSOR
16
INDICATOR
14
22
BATTERY
A/D
13
12
10

FIG. 3

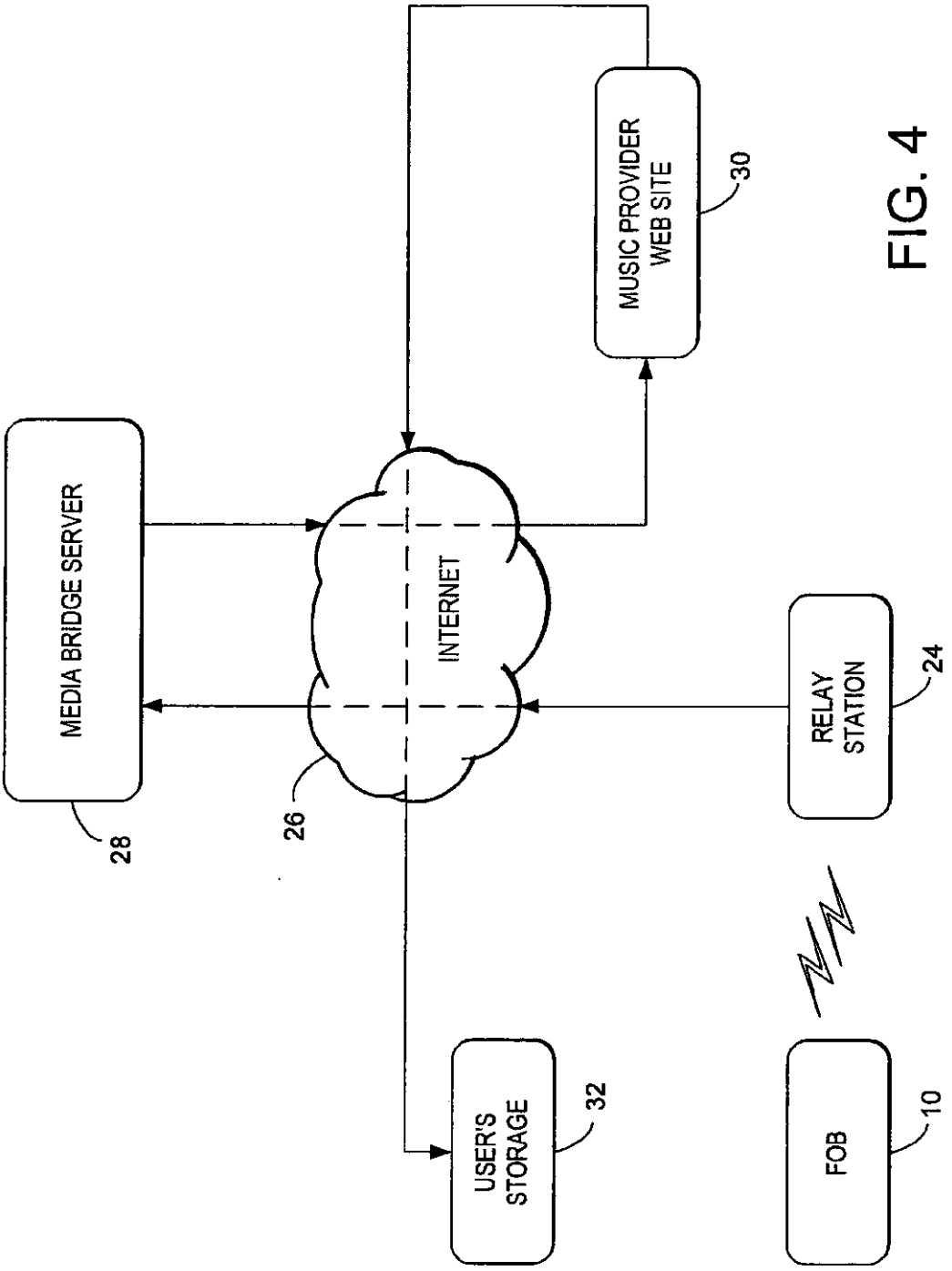

MUSIC PROVIDER WEB SITE
30
MEDIA BRIDGE SERVER
28
INTERNET
26
RELAY STATION
24
USER'S STORAGE
32
FOB
10

FIG. 4

METHODS AND DEVICES RESPONSIVE TO AMBIENT AUDIO

RELATED APPLICATION DATA

This application is a division of application Ser. No. 09/476,686, filed Dec. 30, 1999, which is a continuation-in-part of application 60/134,782, filed May 19, 1999, attached hereto as Appendix A.

The technology detailed in the present application is also related to that detailed in application Ser. No. 09/343,104, filed Jun. 29, 1999 (now abandoned in favor of continuation application Ser. No. 10/764,430, filed Jan. 23, 2004); Ser. No. 09/292,569, filed Apr. 15, 1999 (now abandoned in favor of continuation application Ser. No. 10/379,393, filed Mar. 3, 2003); Ser. No. 09/314,648, filed May 19, 1999 (now U.S. Pat. No. 6,681,028); 60/141,763, filed Jun. 30, 1999; 60/158,015, filed Oct. 6, 1999; 60/163,332, filed Nov. 3, 1999; 60/164,619, filed Nov. 10, 1999; Ser. No. 09/452,023, filed Nov. 30, 1999 (now U.S. Pat. No. 6,408,082); Ser. No. 09/452,021, filed Nov. 30, 1999; and in U.S. Pat. No. 5,862,260.

INTRODUCTION

16 year old Bob struts into the coffee shop down from high school with a couple of buddies, a subtle deep pound in the ambient sound track lets them know they're in the right place. The three of them instinctually pull out of their pockets their audio Birddawgs (a small hand held unit about the size and style of an auto-door-alarm device, or "fob"), and when they see the tiny green light, they smile, high five, and push the big "GoFetch" button in synchrony. That tune will now be waiting for them at home, safely part of their preferred collection and ever-so-thankfully not lost to their collective bad memory (if they even knew the name of the artist and tune title in the first place!).

33 year old Mary is at home listening to the latest batch of holiday tunes being offered up over her 2-decade-long favorite radio station. She's spent many days now half-consciously culling the tunes for that perfect arrangement for the new year's bash that she regrettably agreed to host. 10:40 AM rolls around and some new tune catches her ear, a tune she knows can work well following the jingle-cats rendition of Strawberry Fields. She half jogs over to the stereo and hit[illegible] "GoFetch" button. In a few days, she'll sit down at the computer and put together the final sound track for the gala evening ahead, her play list dutifully waiting for her shuffling instructions and desired start time.

49 year old Jack (the financial analyst) is thoroughly bored sitting in the crowded gate D23 at Dulles. Droning 20 feet up and over his head is the airport network station, currently broadcasting the national weather report. As the segue to the business segment approaches, the teaser review mentions that they'll be having a report on today's rally in the bond market and the driving forces behind it. Jack pulls out his Birddawg-enabled Palm Pilot on the off-chance they actually will have a little depth in the reporting. Indeed, as the segment plays and starts discussing the convoluted effects of Greenspan's speech to the Internet-B-Free society, he taps the "GoFetch" button, knowing that once he gets back to his main browsing environment he will be able to follow dozens of links that the airport network has pre-assigned to the segment.

The foregoing and other features and advantages of the present invention will be more readily apparent from the following detailed description, which proceeds with reference to the accompanying figures.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** is a diagram showing the participants, and channels, involved in the distribution of music.

FIG. **2** shows a conceptual model of how music artists, record labels, and E-Music distributors can all interact with a Media Asset Management System.

FIG. **3** is a block diagram of a device according to one embodiment of the present invention.

FIG. **4** is a block diagram of a system in which the device of FIG. **3** may be utilized.

DETAILED DESCRIPTION

Referring to FIG. **3**, a device **10** according to one embodiment of the present invention includes a microphone **12**, an A/D converter **13**, a processor **14**, one or more indicators **16**, one or more buttons **18**, a wireless interface **20**, and a power source **22**.

The device can be packaged in a small plastic housing, preferably as small as is practical (e.g., sized and configured to serve as a key chain ornament, perhaps akin to the Tomagatchi toys that were recently popular). The housing has one or more small holes to permit audio penetration through the housing to the microphone **12**.

The processor **14** can take various forms, including a dedicated hardware device (e.g., an ASIC), a general purpose processor programmed in accordance with instructions stored in non-volatile RAM memory, etc.

The indicators **16** can be as simple as a single LED lamp, or as complex as an alphanumeric LCD or other multi-element display. In one embodiment, the indicator simply indicates when the processor has decoded a watermark in audio sensed by the microphone. More elaborate signaling techniques can of course be used, including two- or three-color LEDs that can be used to signal different states with different colors, indicators with flashing patterns or changing displays, etc.

The buttons **18** are used by the user to indicate an interest in the audio just-heard. In one embodiment, there is a single button **18**, and it is emblazoned with a stylized legend that can serve as a trademark or service mark, e.g., GetIt!, GoFetch, Birddawg, something Batman-esque ("Wham," "Zappp," "Pow!!," etc.), or something more mundane (e.g., Capture).

The power source **22** can be a battery, solar cel[illegible]age capacitor, or other source of energy suitable for powering the components of the device **10**.

The wireless interface **20** serves to exchange data with a relay station **24** (FIG. **4**). In one embodiment, the interface is radio-based, and provides a one-way communications channel. In other embodiments other wireless technologies can be used (e.g., IR), and/or two-way communication can be provided.

The relay station can be a cellular repeater (if the interface transmits using cellular frequencies and protocols), or a local receiver, e.g., associated with the user's computer. The relay station can also be a paging system relay station (e.g., as are used for two-way pagers), or may be a low earth orbit satellite-based repeater.

In operation, the processor monitors the ambient audio for the presence of encoded data, e.g., a digital watermark, and decodes same. If power considerations permit, the device is "always-on." In other embodiments, one of the buttons **18** can be used to awaken the device. In such other embodiments, another button-press can serve to turn-off the device, or the device can power-down after a predetermined period, e.g., of not sensing any watermarked audio.

A number of techniques for watermarking audio (and decoding same) are known, as illustrated by U.S. Pat. Nos. 5,862,260, 5,963,909, 5,940,429, 5,940,135, 5,937,000, 5,889,868, 5,833,432, 5,945,932, WO9939344 (corresponding to U.S. application Ser. No. 09/017,145, now U.S. Pat. No. 6,145,081), and WO9853565 (corresponding to U.S. application Ser. Nos. 08/858,562 and 08/974,920, now U.S. Pat. Nos. 5,940,135 and 6,175,627, respectively). Commercially-available audio watermarking software includes that available from AudioTrack, Verance (formerly Aris/Solana), Cognicity, Liquid Audio, and others.

The data payload encoded by the watermark (the audio-ID) may take various forms. One is a Digital Object Identifier—an ID corresponding to the standardized digital object numbering system promulgated by the International DOI Foundation (www.doi.org). Another is to include plural data fields variously representing, e.g., the name of the publisher, the name of the artist, the title of the work, the date of publication, etc., etc. Another is to encode a unique identifier (UID), e.g., of 16-64 bits. The UID serves as an index to a remote database where additional information (e.g., publisher, artist, title, date of publication, etc., are stored). The data transmitted from the device **10** to the relay station **24** typically includes some or all of the watermark payload data, and also includes data identifying the device **10**, or its user (user-ID data). Again, this data can include several data fields (e.g. user name, audio delivery information such as email address or URL, age, gender, model of device **10**, etc.). Alternatively, a serial number or other unique identifier can be used, which serves as an index to a database have a corresponding record of information relating to the user and/or device.

The audio-ID and user-ID data are typically formatted and encoded by the device **10** according to a protocol that provides error correcting, framing, and other data useful in assuring reliable transmission to the relay station, and/or for further transport.

Some embodiments of device **10** recognize just a single form of watermarking, and can understand only payload data presented in a single format. In other embodiments, the device may be capable of recognizing watermarking according to several different techniques, and with several different payload formats. This latter functionality can be achieved, e.g., by cyclically trying different decoding techniques until one that produces valid output data (e.g., by reference to a checksum or other indicia) is obtained. That decoding technique and payload interpretation can thereafter be used until valid output data is no longer obtained.

In some embodiments, the device **10** transmits data to the relay station at the moment the user presses the button **18**. In other embodiments, a store-and-forward mode is used. That is, when the user presses the button **18**, the decoded watermark data is stored in memory within the device. Thereafter, e.g., when the device is coupled with a "nest" or "holster" at the user's computer (or when download capability is otherwise activated), the stored data is downloaded—either through that device or otherwise.

The infrastructure between the device **10** and delivery of the audio to its ultimate destination can take myriad forms. One is shown in FIG. **4**. In this arrangement, some or all of the data received by the relay station **24** is routed through the internet **26** to a server **28**. (The server **28** can be a "MediaBridge" server of the type described, e.g., in the assignee's application 60/164,619, filed Nov. 10, 1999, and Ser. No. 09/343,104, filed Jun. 29, 1999.) Server **28** parses the data and routes some or all of it to a data repository **30** at which the audio requested by the user is stored. This repository, in turn, dispatches the audio to the user (e.g., to a computer, media player, storage device, etc.), again through the internet. (Address information detailing the destination **32** of the audio may be included in the data sent from the device **10**, or can be retrieved from a database at the server **28** based on a user-ID sent from the device **10**.)

In some embodiments, the repository **30** (which may be co-located with server **28**, or not) includes various data beyond the audio itself. For example, the repository can store a collection of metadata (e.g., XML tags) corresponding with each stored item of audio. This metadata can be transmitted to the user's destination **32**, or can be used, e.g., for rights management purposes (to limit the user's reproduction or re-distribution rights for the audio, etc.), to establish a fee for the audio, etc. One suitable metatag standard is that under development by <indecs> (Interoperability of Data in E-Commerce Systems, www.indecs.org).

The audio data can be delivered in streaming form, such as using technology available from RealNetworks (RealAudio), Microsoft (Windows Media Player), MP3, Audiobase, Beatnik, Bluestreak.com, etc. The former three systems require large (e.g., megabytes) player software on the receiving (client) computer; the latter do not but instead rely, e.g., on small Java applets that can be downloaded with the music.

Alternatively, the audio can be delivered in a file format. In some embodiments the file itself is delivered to the user's destination **32** (e.g., as an email attachment). In others, the user is provided a URL to permit access to, or downloading of, the audio. (The URL may be a web site that provides an interface through which the user can pay for the requested music, if pre-payment hasn't been arranged.)

The user's destination **32** is typically the user's own computer. If a "live" IP address is known for that computer (e.g., by reference to a user profile database record stored on the server **28**), the music can be transferred immediately. If the user's computer is only occasionally connected to the internet, the music can be stored at a web site (e.g. protected with a user-set password), and can be downloaded to the user's computer whenever it is convenient.

In other embodiments, the destination **32** is a personal music library associated with the user. The library can take the form, e.g., of a hard-disk or semiconductor memory array in which the user customarily stores music. This storage device is adapted to provide music data to one or more playback units employed by the user (e.g. a personal MP3 player, a home stereo system, a car stereo system, etc.). In most installations, the library is physically located at the user's residence, but could be remotely sited, e.g. consolidated with the music libraries of many other users at a central location.

The personal music library can have its own internet connection. Or it can be equipped with wireless capabilities, permitting it to receive digital music from wireless broadcasts (e.g. from a transmitter associated with the server **28**). In either case, the library can provide music to the user's playback devices by short-range wireless broadcast.

In many embodiments, technology such as that available from Sonicbox, permits audio data delivered to the computer to be short range FM-broadcast by the user's computer to nearby FM radios using otherwise-unused radio spectrum.

Some implementations of the present invention support several different delivery technologies (e.g., streaming, file, URL), and select among them in accordance with the profiles of different users.

Payment for the audio (if needed) can be accomplished by numerous means. One is by charging of a credit card account associated with the user (e.g., in a database record corresponding to the user-ID).

Some implementations of the invention make use of secure delivery mechanisms, such as those provided by InterTrust, Preview Systems, etc. In addition to providing secure containers by which the audio is distributed, such systems also include their own secure payment facilities.

By such arrangements, a user can conveniently compile an archive of favorite music—even while away from home.

To provide a comprehensive disclosure without unduly lengthening this specification, the disclosures of the applications and patents cited above are incorporated herein by reference.

Having described and illustrated the principles of our invention with reference to a preferred embodiment and several variations thereof, it should be apparent that the detailed embodiment is illustrative only and should not be taken as limiting the scope of our invention.

For example, while the invention is illustrated with reference to a button that is activated by the user to initiate capture of an audio selection, other interfaces can be used. For example, in some embodiments it can be a voice-recognition system that responds to spoken commands, such as "capture" or "record." Or it can be a form of gesture interface.

Likewise, while the invention is illustrated with reference to a stand-alone device, the same functionality can be built-into radios (including internet-based radios that receive wireless IP broadcasts), computer audio systems, and other appliances. In such case the microphone can be omitted and, in some cases, the wireless interface as well. (The data output from the device can be conveyed, e.g., through the network connection of an associated computer, etc.)

Moreover, while the invention is illustrated with reference to an embodiment in which audio, alone, is provided to the user, this need not be the case. As in the Dulles airport scenario in the introduction, the server **28** can provide to the user several internet links associated with the sensed audio. Some of these links can provide commerce opportunities (e.g., to purchase a CD on which the sensed audio is recorded). Others can direct the user to news sites, concert schedules, fan-club info, etc. In some such embodiments, the ancillary information is provided to the user without the audio itself.

Although not particularly detailed, the data provided to the user's destination typically includes information about the context in which the data was requested. In a simple case this can be the time and date on which the user pressed the Capture button. Other context information can be the identification of other Birddawg devices **10** that were nearby when the Capture button was pressed. (Such information can be gleaned, e.g., by each device transmitting a brief WhoAmI message periodically, receiving such messages from other nearby devices, and logging the data thus received.)

Still other context information might be the location from which the Capture operation was initiated. This can be achieved by decoding of a second watermark signal, e.g., on a low level white-noise broadcast. The public address system in public places, for example, can broadcast a generally-indiscernable noise signal that encodes a watermark signal. Devices **10** can be arranged to detect two (or more) watermarks from the same audio stream, e.g., by reference to two pseudo-random sequences with which the different watermarks are encoded. One identifies the audible audio, the other identifies the location. By such an arrangement, for example, the device **10** can indicate to the server **28** (and thence to the user destination **32**) the location at which the user encountered the audio. (This notion of providing location context information by subliminal audio that identifies the location has powerful applications beyond the particular scenario contemplated herein.)

In some embodiments, the device **10** can buffer watermark information from several previous audio events, permitting the user to scroll back and select (e.g., in conjunction with a screen display **16**) the ID of the desired audio.

An arrangement like the foregoing may require that the decoded watermark information be interpreted for the user, so that the user is not presented simply a raw binary watermark payload. The interpreted information presented to the user can comprise, e.g., the source (CNN Airport News, WABC Radio, CD-ROM, MTV), the artist (Celine Dion), the title (That's the Way It Is), and/or the time decoded (3:38:02 p.m.), etc.

One way to achieve the foregoing functionality is to convey both the binary UID payload and abbreviated text (e.g., 5- or 6-bit encoded) through the watermark "channel" on the audio. In one such arrangement, the watermark channel conveys data a UID, four characters of text, and associated error-correcting bits, every ten seconds. In the following ten seconds the same UID is conveyed, together with the next four characters of text.

Another way to achieve such functionality is to provide a memory in the device **10** that associates the watermark payload (whether UID or field-based) with corresponding textual-data (e.g., the source/artist/title referenced above). A 1 megabyte semiconductor non-volatile RAM memory, for example, can serve as a look-up table, matching code numbers to artist names and song titles. When the user queries the device to learn the identify of a song (e.g., by operating a button **18**), the memory is indexed in accordance with one or more fields from the decoded watermark, and the resulting textual data from the memory (e.g. source/artist/title) is presented to the user.

Such a memory will commonly require periodic updating. The wireless interface **20** in device **10** can include reception capabilities, providing a ready mechanism for providing such updated data. In one embodiment, the device "awakens" briefly at otherwise idle moments and tunes to a predetermined frequency at which updated data for the memory is broadcast, either in a baseband broadcast channel, or in an ancillary (e.g. SCA) channel.

In variants of the foregoing, internet delivery of update data for the memory can be substituted for wireless delivery. For example, a source/artist/title memory in the device **10** can be updated by placing the device in a "nest" every evening. The nest (which may be integrated with a battery charger for the appliance) can have an internet connection, and can exchange data with the device by infrared, inductive, or other proximity-coupling technologies, or through metal contacts. Each evening, the nest can receive an updated collection of source/artist/title data, and can re-write the memory in the device accordingly. By such arrangement, the watermark data can always be properly interpreted for presentation to the user.

The "Capture" concepts noted above can be extended to other functions as well. One is akin to forwarding of email. If a consumer hears a song that another friend would enjoy, the listener may send a copy of the song to the friend. This instruction can be issued by pressing a "Send" button, or by invoking a similar function on a graphical (or voice- or gesture-responsive) user interface. In response, the device so-instructed can query the person as to the recipient. The person can designate the desired recipient(s) by scrolling through a pre-stored list of recipients to select the desired one. (The list can be entered through a computer to which the device is coupled.) Alternatively, the user can type-in a name (if the device provides a keypad), or a portion thereof sufficient to uniquely identify the recipient. Or the person may speak the recipient's name. As is conventional with hands-free vehicle

cell phones, a voice recognition unit can listen to the spoken instructions and identify the desired recipient. An "address book"-like feature has the requisite information for the recipient (e.g., the web site, IP address, or other data identifying the location to which music for that recipient should stored or queued, the format in which the music should be delivered, etc.) stored therein. In response to such command, the appliance dispatches instructions to the server **28**, including an authorization to incur any necessary charges (e.g., by debiting the sender's credit card). Again, the server **28** attends to delivery of the music in a desired manner to the specified recipient.

Still further, a listener may query the device (by voice, GUI or physical button, textual, gesture, or other input) to identify CDs on which the ambient audio is recorded. Or the listener may query the device for the then-playing artist's concert schedule. Again, the appliance can contact a remote database and relay the query, together with the user ID and audio ID data. The database locates the requested data, and presents same to the user—either through a UI on device **10**, or to the destination **32**. If desired, the user can continue the dialog with a further instruction, e.g., to buy one of the CDs on which the then-playing song is included. Again, this instruction may be entered by voice, GUI, etc., and dispatched from the device to the server, which can then complete the transaction in accordance with pre-stored information (e.g. credit card account number, mailing address, etc.). A confirming message can be relayed to the device **10** or destination **32** for presentation to the user.

While the invention particularly contemplates audio, the principles detailed above find applications in many other media, and in many other applications of the MediaBridge server **28**.

Moreover, while the invention particularly contemplates watermarks as the channel by which audio is identified, in other embodiments different techniques can be used. For example, digital radio protocols provide ID fields by which audio can be identified. Similarly, IP protocols for internet delivery of radio include identification fields within their packet formats. Accordingly, audio distributed according to formats that include audio IDs therein can likewise be employed according to the present invention.

In view of the many embodiments to which the principles of our invention may be applied, it should be apparent that the detailed embodiment is illustrative only and should not be taken as limiting the scope of the invention. Rather, I claim as our invention all such modifications as may fall within the scope and spirit of the following claims, and equivalents thereto.

We claim:

1. A device comprising a housing sized for carrying in a user's pocket and including:

a transducer to receive ambient audio and to output electrical signals corresponding thereto;

a watermark detector system coupled to the transducer and adapted to produce payload information by extracting a digital watermark embedded in the electrical signals corresponding to the ambient audio;

a memory storing first identification information; and

an interface that receives at least some of both the decoded payload information and the first identification information for transmission from said device;

wherein the watermark detector system is adapted to decode watermarks encoded by plural different techniques.

2. The device of claim **1** wherein the watermark detector is adapted to decode watermarks having plural different payload formats.

3. The device of claim **1** wherein the watermark detector is adapted to cyclically apply different watermark decoding techniques, and check for valid output data.

4. The device of claim **1** wherein the watermark detector system provides data from two different decoded payloads to the interface, for transmission from the device.

5. A device comprising a housing sized for carrying in a user's pocket and including:

a transducer to receive ambient audio and to output electrical signals corresponding thereto;

a processor coupled to the transducer and adapted to transform the electrical signals corresponding to the ambient audio, to produce transformed data based thereon;

a memory storing first identification information; and

an interface that receives at least some of both the transformed data and the first identification information for transmission to a remote station;

wherein the first identification information comprises information identifying a model of the device.

6. The device of claim **5** further comprising a graphical user interface displayed on a screen, the graphical user interface including a control by which a user can signal interest in the ambient audio.

7. The device of claim **5** further comprising a gesture interface by which a user can signal interest in the ambient audio.

8. The device of claim **5** including a data store that stores information about audio previously encountered by the user, and a user interface operable by the user to scroll back through the stored information and make a selection therefrom.

9. The device of claim **5** wherein the processor comprises a digital watermark decoder.

10. A method employing a device sized for carrying in a user's pocket, the method comprising the acts:

through a user interface of the device, receiving a user input signaling interest in ambient audio;

capturing the ambient audio with a microphone;

converting the captured ambient audio into digital form;

providing the digital form of the ambient audio to a processor;

receiving data corresponding to the ambient audio back from the processor;

recalling first identification information from a memory, the first identification information comprising information identifying a model of the device;

transmitting both the received data and the recalled first identification information to a remote computer, together with information relating to a location at which the ambient audio was received;

receiving information relating to the ambient audio in response; and

controlling an operational aspect of the device in accordance with the received information.

11. The method of claim **10** that additionally includes, at the processor, deriving said data from the digital form, and transmitting the data to the device.

12. The method of claim **11** wherein the deriving comprises applying a digital watermark decoding procedure to the digital form.

13. A method employing a device sized for carrying in a user's pocket, the method comprising the acts:

through a user interface of the device, receiving a user input signaling interest in ambient audio;

capturing the ambient audio with a microphone;

converting the captured ambient audio into digital form;

providing the digital form of the ambient audio to a processor;

receiving data corresponding to the ambient audio back from the processor;

recalling first identification information from a memory, the first identification information comprising information identifying a model of the device;

transmitting both the received data and the recalled first identification information to a remote computer, together with context information including information identifying a further device;

receiving information relating to the ambient audio in response; and

controlling an operational aspect of the device in accordance with the received information.

14. The method of claim **13** wherein the transmitted context information includes information identifying a further device proximate to the user device.

15. The method of claim **13** that additionally includes, at the processor, deriving said data from the digital form, and transmitting the data to the device.

16. The method of claim **15** wherein the deriving comprises applying a digital watermark decoding procedure to the digital form.

17. A method employing a device sized for carrying in a user's pocket, the method comprising the acts:

through a user interface of the device, receiving a user input signaling interest in ambient audio;

capturing the ambient audio with a microphone;

converting the captured ambient audio into digital form;

providing the digital form of the ambient audio to a processor;

receiving data corresponding to the ambient audio back from the processor;

recalling first identification information from a memory, the first identification information comprising information identifying a model of the device;

transmitting both the received data and the recalled first identification information to a remote computer;

receiving information relating to the ambient audio in response; and

controlling an operational aspect of the device in accordance with the received information;

wherein the device includes a data store in which information relating to previously encountered audio is stored, and the method includes scrolling back and selecting from said stored information.

18. The method of claim **17** that additionally includes, at the processor, deriving said data from the digital form, and transmitting the data to the device.

19. The method of claim **18** wherein the deriving comprises applying a digital watermark decoding procedure to the digital form.

* * * * *

US007590259B2

(12) **United States Patent**
Levy et al.

(10) **Patent No.: US 7,590,259 B2**
(45) **Date of Patent: *Sep. 15, 2009**

(54) **DERIVING ATTRIBUTES FROM IMAGES, AUDIO OR VIDEO TO OBTAIN METADATA**

(75) Inventors: **Kenneth L. Levy**, Stevenson, WA (US); **Geoffrey B. Rhoads**, West Linn, OR (US)

(73) Assignee: **Digimarc Corporation**, Beaverton, OR (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **11/927,075**

(22) Filed: **Oct. 29, 2007**

(65) **Prior Publication Data**

US 2008/0139182 A1 Jun. 12, 2008

Related U.S. Application Data

(60) Continuation of application No. 10/338,032, filed on Jan. 6, 2003, now Pat. No. 7,349,552, which is a division of application No. 09/563,664, filed on May 2, 2000, now Pat. No. 6,505,160, which is a continuation-in-part of application No. 09/476,686, filed on Dec. 30, 1999, application No. 11/927,075, which is a continuation of application No. 10/338,031, filed on Jan. 6, 2003, now Pat. No. 7,333,957, which is a division of application No. 09/563,664, filed on May 2, 2000, now Pat. No. 6,505,160, which is a continuation-in-part of application No. 08/746,613, filed on Nov. 12, 1996, now Pat. No. 6,122,403, which is a continuation-in-part of application No. 08/649,419, filed on May 16, 1996, now Pat. No. 5,862,260, and a continuation-in-part of application No. PCT/US96/06618, filed on May 7, 1996, and a continuation-in-part of application No. 08/508,083, filed on Jul. 27, 1995, now Pat. No. 5,841,978.

(51) **Int. Cl.**
G06K 9/00 (2006.01)

(52) **U.S. Cl.** **382/100**; 382/313

(58) **Field of Classification Search** 382/100, 382/232, 313, 317; 713/176; 380/210, 287, 380/54; 455/3.01, 3.06; 725/62
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

3,406,344 A 10/1968 Hopper 325/50

(Continued)

FOREIGN PATENT DOCUMENTS

EP 493 091 A1 7/1992

(Continued)

OTHER PUBLICATIONS

Foote, "An Overview of Audio Information Retrieval," Multimedia Systems, v.7 n. 1, p. 2-10, Jan. 1999.

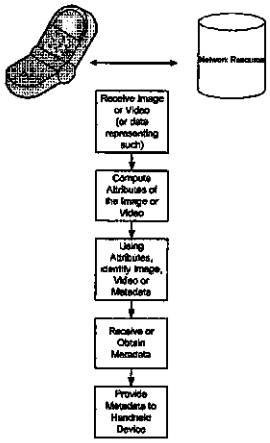
(Continued)

Primary Examiner—Andrew W Johns

(57) **ABSTRACT**

The present invention relates generally to obtaining metadata associated with images, audio and video. Once claim recites a method including: obtaining data corresponding to media content from a handheld device, the data representing picture elements of an image or video or representing audible portions of an audio signal; computing attributes of the data using a processor, said act of computing utilizes the processor to operate on the data; using computed attributes of the data to identify the media content or to identify metadata associated with the media content; obtaining metadata associated with the media content; and providing the metadata to the handheld device from a ne[illegible] resource. Other combinations and claims are provided as well.

64 Claims, 3 Drawing Sheets



U.S. PATENT DOCUMENTS

3,586,781	A	6/1971	Jones	179/15
3,898,390	A	8/1975	Wells et al.	179/41 A
3,919,479	A	11/1975	Moon et al.	179/1 SB
3,950,782	A	4/1976	Carey et al.	360/72
3,984,624	A	10/1976	Waggener	178/5.6
4,025,851	A	5/1977	Haselwood et al.	325/31
4,230,990	A	10/1980	Lert, Jr. et al.	455/67
4,284,846	A	8/1981	Marley	179/1 SE
4,313,197	A	1/1982	Maxemchuk	370/111
4,450,531	A	5/1984	Kenyon et al.	364/604
4,542,525	A	9/1985	Hopf	381/56
4,639,779	A	1/1987	Greenberg	358/142
4,677,466	A	6/1987	Lert, Jr. et al.	358/84
4,697,209	A	9/1987	Kiewit et al.	358/84
4,739,398	A	4/1988	Thomas et al.	358/84
4,796,292	A	1/1989	Thomas	379/91
4,805,020	A	2/1989	Greenberg	358/147
4,807,031	A	2/1989	Broughton et al.	358/142
4,843,562	A	6/1989	Kenyon et al.	364/487
4,858,000	A	8/1989	Lu	358/84
4,907,264	A	3/1990	Seiler et al.	379/355
4,931,871	A	6/1990	Kramer	358/142
4,945,412	A	7/1990	Kramer	358/142
4,967,273	A	10/1990	Greenberg	358/142
4,969,041	A	11/1990	O'Grady et al.	358/142
4,972,471	A	11/1990	Gross et al.	380/3
5,001,696	A	3/1991	Baldwin	364/521
5,019,899	A	5/1991	Boles et al.	358/84
5,031,228	A	7/1991	Lu	382/38
5,077,608	A	12/1991	Dubner	358/183
5,103,459	A	4/1992	Gilhousen et al.	375/1
5,115,326	A	5/1992	Burgess et al.	358/440
5,204,902	A	4/1993	Reeds, III et al.	380/23
5,210,820	A	5/1993	Kenyon	395/2
5,315,098	A	5/1994	Tow	235/494
5,319,735	A	6/1994	Preuss et al.	395/2.14
5,331,547	A	7/1994	Laszlo	364/413.01
5,398,336	A	3/1995	Tantry et al.	395/600
5,415,553	A	5/1995	Szmidla	434/309
5,420,943	A	5/1995	Mak	382/313
5,436,653	A	7/1995	Ellis et al.	348/2
5,444,230	A	8/1995	Baldwin et al.	235/462
5,444,779	A	8/1995	Daniele	380/3
5,473,631	A	12/1995	Moses	375/202
5,474,457	A	12/1995	Bromley	434/311
5,480,306	A	1/1996	Liu	434/156
5,[illegible]	A	1/1996	Thomas et al.	348/1
5,486,686	A	1/1996	Zdybel, Jr. et al.	235/375
5,509,074	A	4/1996	Choudhury et al.	380/23
5,530,852	A	6/1996	Meske, Jr. et al.	395/600
5,572,653	A	11/1996	DeTemple et al.	395/501
5,574,519	A	11/1996	Manico et al.	396/429
5,574,962	A	11/1996	Fardeau et al.	455/2
5,577,266	A	11/1996	Takahisa et al.	455/66
5,579,124	A	11/1996	Aijala et al.	386/96
5,581,800	A	12/1996	Fardeau et al.	455/2
5,584,070	A	12/1996	Harris et al.	455/346
5,598,557	A	1/1997	Doner et al.	395/605
5,606,609	A	2/1997	Houser et al.	380/4
5,606,668	A	2/1997	Shwed	395/200.11
5,613,004	A	3/1997	Cooperman et al.	382/28
5,621,800	A	4/1997	Weng et al.	380/49
5,640,193	A	6/1997	Wellner	348/7
5,671,282	A	9/1997	Wolff et al.	380/25
5,687,191	A	11/1997	Lee et al.	375/216
5,703,795	A	12/1997	Mankovitz	364/514 R
5,708,478	A	1/1998	Tognazzini	348/552
5,710,834	A	1/1998	Rhoads	382/232
5,737,025	A	4/1998	Dougherty et al.	348/473
5,740,244	A	4/1998	Indeck et al.	380/4
5,745,569	A	4/1998	Moskowitz et al.	380/4
5,745,782	A	4/1998	Conway	395/806
5,751,854	A	5/1998	Saitoh et al.	382/218
5,754,981	A	5/1998	Veeneman et al.	705/26
5,761,606	A	6/1998	Wolzien	455/6.2
5,764,763	A	6/1998	Jensen et al.	380/6
5,765,152	A	6/1998	Erickson	707/9
5,774,452	A	6/1998	Wolosewicz	370/212
5,781,914	A	7/1998	Stork et al.	707/506
5,782,692	A	7/1998	Stelovsky	463/1
5,790,172	A	8/1998	Imanaka	348/7
5,801,689	A	9/1998	Huntsman	345/329
5,801,702	A	9/1998	Dolan et al.	345/357
5,809,160	A	9/1998	Powell et al.	382/100
5,822,360	A	10/1998	Lee et al.	375/200
5,822,432	A	10/1998	Moskowitz et al.	380/28
5,822,436	A	10/1998	Rhoads	380/54
5,832,119	A	11/1998	Rhoads	382/232
5,838,458	A	11/1998	Tsai	358/402
5,841,886	A	11/1998	Rhoads	382/115
5,841,978	A	11/1998	Rhoads	395/200.47
5,842,162	A	11/1998	Fineberg	704/233
5,848,413	A	12/1998	Wolff	707/10
5,862,260	A	1/1999	Rhoads	382/232
5,869,819	A	2/1999	Knowles et al.	235/375
5,870,552	A	2/1999	Dozier et al.	395/200.49
5,872,531	A	2/1999	Johnson et al.	341/110
5,901,224	A	5/1999	Hecht	380/4
5,905,248	A *	5/1999	Russell et al.	235/462
5,918,223	A	6/1999	Blum et al.	707/1
5,932,863	A	8/1999	Rathus	235/462.15
5,937,000	A	8/1999	Lee et al.	375/200
5,938,727	A	8/1999	Ikeda	709/218
5,963,957	A	10/1999	Hoffberg	707/104
5,969,283	A	10/1999	Looney et al.	84/609
5,978,773	A	11/1999	Hudetz et al.	705/23
5,978,791	A	11/1999	Farber et al.	707/2
5,982,956	A	11/1999	Lahmi	382/306
5,986,651	A	11/1999	Reber et al.	345/335
5,988,897	A	11/1999	Pierce et al.	400/61
5,991,500	A	11/1999	Kanota et al.	386/94
5,991,737	A	11/1999	Chen	705/26
5,995,105	A	11/1999	Reber et al.	345/356
5,999,569	A	12/1999	Oshima	375/265
6,002,443	A	12/1999	Iggulden	348/553
6,005,501	A	12/1999	Wolosewicz	341/52
6,064,737	A	5/2000	Rhoads	380/23
6,076,734	A	6/2000	Dougherty et al.	235/462.01
6,081,827	A	6/2000	Reber et al.	709/200
6,081,830	A	6/2000	Schindler	709/204
6,088,455	A	7/2000	Logan et al.	380/200
6,098,106	A	8/2000	Philyaw et al.	709/238
6,121,530	A	9/2000	Sonoda	84/609
6,122,403	A	9/2000	Rhoads	382/233
6,122,526	A	9/2000	Parulski et al.	455/556
6,125,172	A	9/2000	August et al.	379/110.01
6,147,940	A	11/2000	Yankowski	369/30
6,148,333	A	11/2000	Guedalia et al.	709/219
6,151,624	A	11/2000	Teare et al.	709/217
6,169,541	B1	1/2001	Smith	345/327
6,181,817	B1	1/2001	Zabih et al.	382/170
6,185,316	B1	2/2001	Buffam	382/115
6,188,010	B1	2/2001	Iwamura	84/609
6,199,048	B1	3/2001	Hudetz et al.	705/23
6,222,807	B1	4/2001	Min-Jae	369/58
6,243,480	B1	6/2001	Zhao et al.	382/100
6,278,781	B1	8/2001	Rhoads	380/247
6,282,362	B1	8/2001	Murphy et al.	386/46
6,286,036	B1	9/2001	Rhoads	709/217
6,304,523	B1	10/2001	Jones et al.	369/30
6,311,214	B1	10/2001	Rhoads	709/217
6,317,881	B1	11/2001	Shah-Nazaroff et al.	725/24
6,324,573	B1	11/2001	Rhoads	709/217
6,338,094	B1	1/2002	Scott et al.	709/245

6,345,256	B1	2/2002	Milsted et al.	705/1
6,389,055	B1	5/2002	August et al.	375/130
6,408,331	B1	6/2002	Rhoads	709/217
6,415,280	B1	7/2002	Farber et al.	707/2
6,433,946	B2	8/2002	Ogino	360/60
6,456,725	B1	9/2002	Cox et al.	382/100
6,466,670	B1	10/2002	Tsuria et al.	380/202
6,505,160	B1	1/2003	Levy et al.	704/270
6,510,234	B1	1/2003	Cox et al.	382/100
6,522,770	B1	2/2003	Seder et al.	382/100
6,523,175	B1	2/2003	Chan	725/15
6,526,449	B1	2/2003	Philyaw et al.	709/238
6,539,095	B1	3/2003	Rhoads	381/73.1
6,542,927	B2	4/2003	Rhoads	709/217
6,546,112	B1	4/2003	Rhoads	382/100
6,553,129	B1	4/2003	Rhoads	382/100
6,560,349	B1	5/2003	Rhoads	382/100
6,587,821	B1	7/2003	Rhoads	704/270
6,590,998	B2	7/2003	Rhoads	382/100
6,611,607	B1	8/2003	Davis et al.	382/100
6,640,306	B1	10/2003	Tone et al.	713/201
6,647,130	B2	11/2003	Rhoads	382/100
6,650,761	B1	11/2003	Rodriguez et al.	382/100
6,664,976	B2	12/2003	Lofgren et al.	345/634
6,674,993	B1	1/2004	Tarbouriech	455/2.01
6,675,146	B2	1/2004	Rhoads	704/270
6,681,028	B2	1/2004	Rodriguez et al.	382/100
6,681,029	B1	1/2004	Rhoads	382/100
6,694,042	B2	2/2004	Seder et al.	382/100
6,694,043	B2	2/2004	Seder et al.	382/100
6,697,948	B1	2/2004	Rabin et al.	713/200
6,728,390	B2	4/2004	Rhoads et al.	382/100
6,748,360	B2	6/2004	Pitman et al.	704/270
6,748,533	B1	6/2004	Wu et al.	713/176
6,768,980	B1	7/2004	Meyer et al.	704/500
6,771,885	B1	8/2004	Agnihotri et al.	386/83
6,772,124	B2	8/2004	Hoffberg et al.	704/270.1
6,775,392	B1	8/2004	Rhoads	382/100
6,782,115	B2	8/2004	Decker et al.	382/100
6,804,379	B2	10/2004	Rhoads	382/101
6,807,676	B1	10/2004	Robbins et al.	725/39
6,829,368	B2	12/2004	Meyer et al.	382/100
6,834,308	B1	12/2004	Ikezoye et al.	709/231
6,850,252	B1	2/2005	Hoffberg	345/716
6,856,977	B1	2/2005	Adelsbach et al.	705/57
6,869,023	B2	3/2005	Hawes	235/494
6,904,185	B1	6/2005	Wilkins et al.	382/311
6,917,691	B2	7/2005	Evans et al.	382/100
6,917,724	B2	7/2005	Seder et al.	382/321
6,931,451	B1	8/2005	Logan et al.	709/231
6,941,275	B1	9/2005	Swierczek	705/26
6,965,873	B1	11/2005	Rhoads	705/26
6,968,337	B2	11/2005	Wold	707/100
6,973,669	B2	12/2005	Daniels	725/112
6,985,600	B2	1/2006	Rhoads et al.	382/100
6,987,862	B2	1/2006	Rhoads	382/100
6,990,453	B2	1/2006	Wang et al.	704/270
7,010,144	B1	3/2006	Davis et al.	382/100
7,012,621	B2	3/2006	Crosby et al.	345/619
7,024,016	B2	4/2006	Rhoads et al.	382/100
7,035,427	B2	4/2006	Rhoads	382/100
7,044,395	B1	5/2006	Davis et al.	235/494
7,047,413	B2	5/2006	Yacobi et al.	713/176
7,050,603	B2	5/2006	Rhoads et al.	382/100
7,051,086	B2	5/2006	Rhoads et al.	709/219
7,058,697	B2	6/2006	Rhoads	709/217
7,076,084	B2	7/2006	Davis et al.	382/100
7,095,871	B2	8/2006	Jones et al.	382/100
7,127,744	B2	10/2006	Levy	726/26
7,136,502	B2	11/2006	Rhoads et al.	382/100
7,164,413	B2	1/2007	Davis et al.	345/163
7,171,018	B2	1/2007	Rhoads et al.	382/100
7,174,031	B2	2/2007	Rhoads et al.	382/107
7,174,293	B2	2/2007	Kenyon et al.	704/231
7,185,201	B2	2/2007	Rhoads et al.	713/176
7,190,971	B1	3/2007	Kawamoto	455/556.1
7,194,752	B1	3/2007	Kenyon et al.	725/22
7,209,571	B2	4/2007	Davis et al.	382/100
7,224,995	B2	5/2007	Rhoads	455/550.1
7,248,715	B2	7/2007	Levy	382/100
7,251,475	B2	7/2007	Kawamoto	455/412.3
7,261,612	B1	8/2007	Hannigan et al.	446/175
7,289,643	B2	10/2007	Brunk et al.	382/100
7,302,574	B2	11/2007	Conwell et al.	713/176
7,333,957	B2	2/2008	Levy et al.	705/58
7,349,552	B2 *	3/2008	Levy et al.	382/100
7,359,528	B2	4/2008	Rhoads	382/100
7,362,781	B2	4/2008	Rhoads	370/522
7,362,879	B2	4/2008	Evans et al.	382/100
7,369,676	B2	5/2008	Hein, III	382/100
7,372,976	B2	5/2008	Rhoads et al.	382/100
7,377,421	B2 *	5/2008	Rhoads	235/375
7,437,430	B2	10/2008	Rhoads	709/219
7,450,734	B2	11/2008	Rodriguez et al.	382/100
7,461,136	B2	12/2008	Rhoads	709/217
2001/0008557	A1	7/2001	Stefik et al.	380/202
2002/0044659	A1	4/2002	Ohta	380/241
2002/0178410	A1	11/2002	Haitsma et al.	714/709
2003/0011684	A1	1/2003	Narayanaswami et al.	348/207.99
2003/0040326	A1	2/2003	Levy et al.	455/466
2003/0167173	A1	9/2003	Levy et al.	704/273
2003/0174861	A1	9/2003	Levy et al.	382/100
2004/0128514	A1	7/2004	Rhoads	713/176
2005/0043018	A1	2/2005	Kawamoto	455/414.3
2005/0044561	A1	2/2005	McDonald	725/18
2005/0058319	A1	3/2005	Rhoads et al.	382/100
2005/0091268	A1	4/2005	Meyer et al.	707/103 R
2005/0229107	A1	10/2005	Hull et al.	715/764
2006/0120560	A1	6/2006	Davis et al.	382/100
2006/0174348	A1	8/2006	Rhoads et al.	726/26
2007/0100757	A1	5/2007	Rhoads	705/51
2007/0189533	A1	8/2007	Rhoads	380/247
2007/0195987	A1	8/2007	Rhoads	382/100
2008/0062315	A1	3/2008	Oostveen et al.	348/500
2008/0125083	A1	5/2008	Rhoads	455/410
2008/0273747	A1	11/2008	Rhoads	382/100

FOREIGN PATENT DOCUMENTS

EP	0493091 ...	7/1992
EP	0581317 A2	2/1994
EP	0581317 A2	2/1994
EP	0642060 A2	3/1995
EP	1199878 A2	4/2002
EP	1049320 B1	1/2003
EP	1019868 B1	1/2009
JP	4-335480 A	11/1992
JP	5-037795 A	2/1993
JP	8-050598 A	2/1996
JP	2005-51793 A	2/2005
WO	WO 9400842 A1	1/1994
WO	WO 9510813 A1	4/1995
WO	WO 9514289 A2	5/1995
WO	WO96/36163 A2	11/1996
WO	WO 9702522 A1	1/1997
WO	WO 97/43736 A1	11/1997
WO	WO9743736 A1	11/1997
WO	WO98/03923 A1	1/1998
WO	WO9904568 A1	1/1999
WO	WO0036605 A1	6/2000
WO	WO0079709 A1	12/2000
WO	WO0172030 A2	9/2001
WO	WO0175794 A2	10/2001

WO WO0211123 A2 2/2002
WO WO02082271 A1 10/2002

OTHER PUBLICATIONS

Smith, et al, "Music Information Retrieval Using Audio Input." Proc AAAI Spring Symposium on Intelligent Integration and Use of Text, Image, Video and Audio Corpora, pp. 12-16.
Roy, "Wearable Audio Computer—A Survey of Interaction Techniques," MIT Media Lab, 1997.
Blackburn, "A Tool for Content Based Navigation of Music," ACM Multimedia 98.
deRoure, "Multiagent System for Content Based Navigation of Music," ACM Multimedia, Oct. 99. 4 pp.
Aust, D., "Augmenting Paper Documents with Digital Information in a Mobile Environment," MS Thesis, University of Dortmund, Department of Computer Graphics, Sep. 3, 1996.
Anonymous, Internet-On-A-Disk #7, Dec. 3, 1994.
Anonymous, NL-KR Digest, Oct. 3, 1988.
Arai et al. "Retrieving Electronic Documents with Real-World Objects on InteractiveDESK," UIST '95, Nov. 14, 1995.
Arai, InteractiveDESK: A Computer-Augmented Desk Which Responds to Operations on Real Objects, CHI 95, May 7, 1995.
T. Berners-Lee, L. Masinter, M. McCahill, Uniform Resource Locators (URL), Network Working Group, Request for Comments 1738, Dec. 1994.
T. Berners-Lee. The Original HTTP as defined in 1991.
Aug. 29, 2008 Amendment; Jul. 3, 2008 Office Action in Ex Parte Reexamination; Oct. 16, 2007 Determination— Reexamination Ordered; Jul. 26, 2007 "Request for Ex Party Reexamination"; all from Reexam No. 90/008,779 (Reexam of U.S. Appl. No. 6,199,048).
Jun. 6, 2008 Notice of Allowance; May 3, 2007 Appeal Brief; Aug. 4, 2006 final Office Action; Jun. 15, 2006 Amendment; Feb. 5, 2006 Office Action; Jan. 24, 2006 Amendment; Jul. 13, 2005 final Office Action; all from assignee's U.S. Appl. No. 10/090,775 (issued as 7,437,430).
Hartung et al., Digital Watermarking of Raw and Compressed Video, Proc. SPIE 2952, Digital Compression Technologies and Systems for Video Communications, Oct. 1996, pp. 205-213.
B. Ibrahim, "World-Wide Algorithm Animation", Computer Networks and ISDN Systems, North Holland Publishing, Nov. 1994.
Namba, S. et al., "A Program Identification Code Transmission System Using Low-Frequency Audio Signals," NHK Laboratories Note, Ser. No. 314, Mar. 1985.
Lin, et al., "Generating Robust Digital Signature for Image/Video Authentication," Proc. Multimedia and Se[illegible]y workshop at ACM [illegible]media'98, Sep. 1, 1998, pp. 49-54.
Worring, "Hyperdocument Generation Using OCR and Icon Detection," Proc. 3d Int. Conf. on Doc. Analysis and Recognition, Aug. 14, 1995.
Ghias et al, Query by Humming: Musical Information Retrieval In An Audio Database. In ACM Multimedia, pp. 231-236, Nov. 1995.
Kageyama et al, Melody Retrieval with Humming, Proceedings of Int. Computer Music Conference (ICMC), 1993.
Muscle Fish press release, Muscle Fish's Audio Search Technology to be Encapsulated into Informix Datablade Module, Jul. 10, 1996.
P. Mockapetris, Domain Names- Concepts and Facilities, Network Working Group, Request for Comments 1034, Nov. 1987.
P. Mockapetris, Domain Names- Implementation and Specification, Network Working Group, Request for Comments 1034, Nov. 1987.
Wagner. "Fingerprinting," 1983 IEEE, pp. 18-22.
Wold et al, Content-Based Classification, Search, and Retrieval of Audio, IEEE Multimedia Magazine, Fall, 1996.
Zhao, et al., "Embedding Robust Labels into Images for Copyright Protection," Proceedings of International Congress on Intellectual Property Rights for Specialized Information, Knowledge and New Technology, Aug. 21, 1995.
U.S. Appl. No. 60/000,442, filed Jun. 20, 1995 (Hudetz).
Steele, R. et al., "Embedding Data in Speech using Scrambling Techniques," IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 3, May 1982.
Steele, R. et al., "Simultaneous Transmission of Speech and Data Using Code-Breaking Techniques," The Bell System Tech. Jour., vol. 60, No. 9, pp. 2081-2105, Nov. 1981.
ten Kate, et al., "Digital Audio Carrying Extra Information," Philips Research Labs, pp. 1097-1100, IEEE 1990.
P.C.J. Hill, "Simultaneous Subliminal Signalling in Conventional Sound Circuits," BBC Engineering, No. 90, pp. 14-31, 1972.
Komatsu, N., "Information Security for Facsimile Communication," Journal of the Institute of Image Electronics Engineers of Japan, 1990, vol. 19 No. 4, pp. 229-235.
Hara et al., "An Improved Method of Embedding Data Into Pictures by Modulo Masking," IEEE Transactions on Communications, 1988 vol. COM-36, No. 3, pp. 315-331.
Xydeas et al., "Embedding Data Into Pictures by Modulo Masking," IEEE Transactions on Communications, 1984, vol. COM-32, No. 1, pp. 56-69.
Matsui et al., "Video-Steganography: How to Secretly Embed a Signature in a Picture," IMA Intellectual Property Project Proceedings, Jan. 1994, Vol. 1, Issue 1, pp. 187-205.
Tanaka et al., "A Visual Retrieval System with Private Information for Image Database," International Conference on DSP Applications and Technology, Oct. 1991, pp. 415-421.
Pitas et al., "Applying Signatures on Digital Images," IEEE Workshop on Nonlinear Image and Signal Processing, Neos Marmaras, Greece, pp. 460-463, Jun. 1995.
Szepanski, "A Signal Theoretic Method for Creating Forgery-Proof Documents for Automatic Verification," in Proceedings of the 1979 Carnahan Conference on Crime Countermeasures, University of Kentucky, Lexington, KY, May 16-18, 1979, pp. 101-109.
Itoh, et al., "A Method of Concealed Image for Bi-Level Image," Technical Research Report of the Institute of Electronics, Information and Communication Engineering, Institute of Electronics, Information and Communication Engineering, Japan, vol. 90, No. 152, Aug. 1990, pp. 83-86.
U.S. Appl. No. 09/476,686, filed Dec. 30, 1999 (pending).
U.S. Appl. No. 09/507,096, filed Feb. 17, 2000 (now abandoned).
U.S. Appl. No. 09/574,726, filed May 18, 2000 (pending).
U.S. Appl. No. 09/636,102, filed Aug. 10, 2000 (pending).
U.S. Appl. No. 10/823,997, filed Apr. 13, 2004 (allowed).
Sep. 28, 2001 Examination Report from the European Patent Office; Feb. 6, 2002 Response thereto; May 9, 2000 European Search Report; and claims as originally filed, each from assignee's European patent application No. 00116604.0 (now issued as EP 1049320 B1).
Sep. 18, 2000 Communication (Search Report) from the European Patent Office in Assignee's European Patent Application No. 97 924 752.5 (published as EP1019868).
Nov. 18, 2005 Communication from the European Patent Office in Assignee's European Patent Appli[illegible] No. 97 924 752.5 (published as EP1019868).
Allowed Claims from assignee's Japanese patent application No. 2004-224727 (published as JP 3949679); and Mar. 24, 2006 Notice of Reason(s) For Rejection (English Translation) in the JP 2004-224727 application.
Mar. 31, 2008 Notice of Allowance (including Examiner's search history), and Feb. 21, 2008 Amendment, each from assignee's U.S. Appl. No. 11/265,544 (published as US 2006-0136565 A1).
May 26, 2006 Response to the Nov. 18, 2005 Communication from the EPO in Assignee's European Patent Application No. 97 924 752.5.
Bender et al., "Techniques for data hiding," Proc. SPIE, vol. 2420, pp. 164-173, 1995.
Digimarc, "Frequently Asked Questions About Digimarc Signature Technology," Aug. 1995.
Highwater FBI. "Copyright Protection for Digital Images, Digital Fingerprinting from FBI," Brochure, 4 pp., 1995.
Koch et al. "Copyright Protection for Multimedia Data," Proc. of the Int. Conf. On Digital Media and Electronic Publishing, Leeds, U.K., 15 pages, Dec. 1994.
Johnson, et al., "Bridging the Paper and Electronic Worlds: The Paper User Interface", Interchi '93, pp. 507-512, Apr. 1993.
Whittaker, et al., "Back to the Future: Pen and Paper Technology Supports Complex Group Coordination," CHI '95, Conference on

Human Factors in Computing Systems, Denver, Colorado (May 7-11, 1995) (text copy obtained from ACM).
Peairs, "Iconic Paper," Proceedings of the Third International Conference on Document Analysis and Recognition (ICDAR '95), pp. 1174-1179, 1995.
Newman, William, et al. "A Desk Supporting Computer-Based Interaction with paper Documents," ACM Conference on Human Factors in Computing Systems (CHI '92) May 3-7, 1992, pp. 587-592.
Rao, et al., "Protofoil: Storing and Finding the Information Worker's Paper Documents in an Electronic File Cabinet," Human Factors in Computing Systems (CHI '94), pp. 180-186, Boston, MA, Apr. 1994.
Smith, et al, "Music Information Retrieval Using Audio Input." Proc AAAI Spring Symposium on Intelligent Integration and Use of Text, Image, Video and Audio *Corpora*, pp. 12-16, 1996.
Matsutani. "The Construction of Copyright-Protected Image Data Technical Report of IEICE," ISEC94-58, pp. 59-68, 1995.
Zhao, "A WWW Service to Embed and Prove Digital Copyright Watermarks," Proc. of the European Conference on Multimedia Applications, Services and Techniques, May 1996, 15 pages.
Zhao, et al., "Embedding Robust Labels into Images for Copyright Protection," Proceedings of International Congress on Intellectual Property Rights for Specialized Information, Knowledge and New Technology, Aug. 21, 1995.
Feb. 5, 2009 Notice of Abandonment; Jul. 18, 2009 final Office Action; Jun. 22, 2007 Amendment; Mar. 22, 2007 non-final Office Action; Oct. 31, 2006 Request for Rehearing Decision; Jul. 31, 2006 Request for Rehearing; Mary 31, 2006 Decision on Appeal; May 24, 2005 Reply Brief; Mar. 24, 2005 Examiner's Answer; Aug. 23, 2004 Appeal Brief; Jan. 21, 2004 final Office Action; all from assignee's U.S. Appl. No. 09/636,102.
U.S. Appl. No. 12/258,174, filed Oct. 24, 2008 (including filing receipt, declaration, application data sheet, specification, claims and drawings).
U.S. Appl. No. 12/275,530, filed Nov. 21, 2008 (including filing receipt, declaration, application data sheet, specification, claims and drawings).
Dec. 30, 2008 non-final Office Action; and Oct. 23, 2007 Preliminary Amendment; all from assignee's U.S. Appl. No. 11/331,430 (published as US 2006-0174348 A1).
U.S. Appl. No. 12/275,530, including filing receipt and application data sheet, filed Nov. 21, 2008.
Jan. 29, 2009 Amendment (including application data sheet); Oct. 28, 2008 non-final Office Action; Jun. 30, 2006 Amendment; May 31, 2006 non-final Office Action; and Sep. 21, 2006 final Office Action; all from assignee's U.S. Appl. No. 09/476,686.
Jan. 23, 2009 Amendment submitted with RCE; Dec. 16, 2008 Notice of Allowance; Oct. 20, 2008 Amendment; Jul. 21, 2008 non-final Office Action; all from assignee's U.S. Appl. No. 11/274,758 (published as US 2006-0120560 A1).
Claims, abstract, filing receipt and application data sheet from assignee's U.S. Appl. No. 12/258,174, filed Oct. 24, 2008.
Jul. 28, 2008 Preliminary Amendment from assignee's U.S. Appl. No. 12/116,645, filed May 7, 2008 (published as US 2008-0273747 A1).
Dec. 15, 2008 Notice of Allowance (including Examiner's Amendment); Oct. 3, 2008 Amendment Accompanying RCE (including attachments and declarations); and specification, drawings, claims and abstract as filed on Jan. 6, 2000; all from assignee's U.S. Appl. No. 09/479,304.
Mar. 7, 2007 Notice of Abandonment; Feb. 15, 2007 Letter of Express Abandonment; Aug. 15, 2006 Office Action; and specification, drawings, claims and abstract as filed on Jun. 20, 2002; all from assignee's U.S. Appl. No. 10/177,650 (published as US 2003-0040326 A1).
Sep. 24, 2008 Amendment; Jun. 3, 2008 Office Action; Feb. 28, 2008 Office Action; and specification, drawings, claims and abstract as filed on Feb. 2, 2007; all from assignee's U.S. Appl. No. 11/670,841 (published as US 2007-0189533 A1).
Specification, drawings, claims and abstract as filed on Feb. 2, 2007; all from assignee's U.S. Appl. No. 11/875,551 (published as US 2008-0125083 A1).
May 26, 2006 Response; Nov. 18, 2005 Communication from the European Patent Office; and Sep. 18, 2000 Supplementary Partial European Search Report; all from Assignee's European Patent Application No. 97 924 752.5 (published as EP1019868).
Jun. 1, 2007 Summons to attend oral proceedings from the European Patent Office in Assignee's European Patent Application No. 97 924 752.5 (published as EP1019868), and issuing EP patent No. EP 1019868 B1.
Highwater FBI Ltd., "FBI Presentation: Image Copyright Protection Software," Jul. 1995.

* cited by examiner

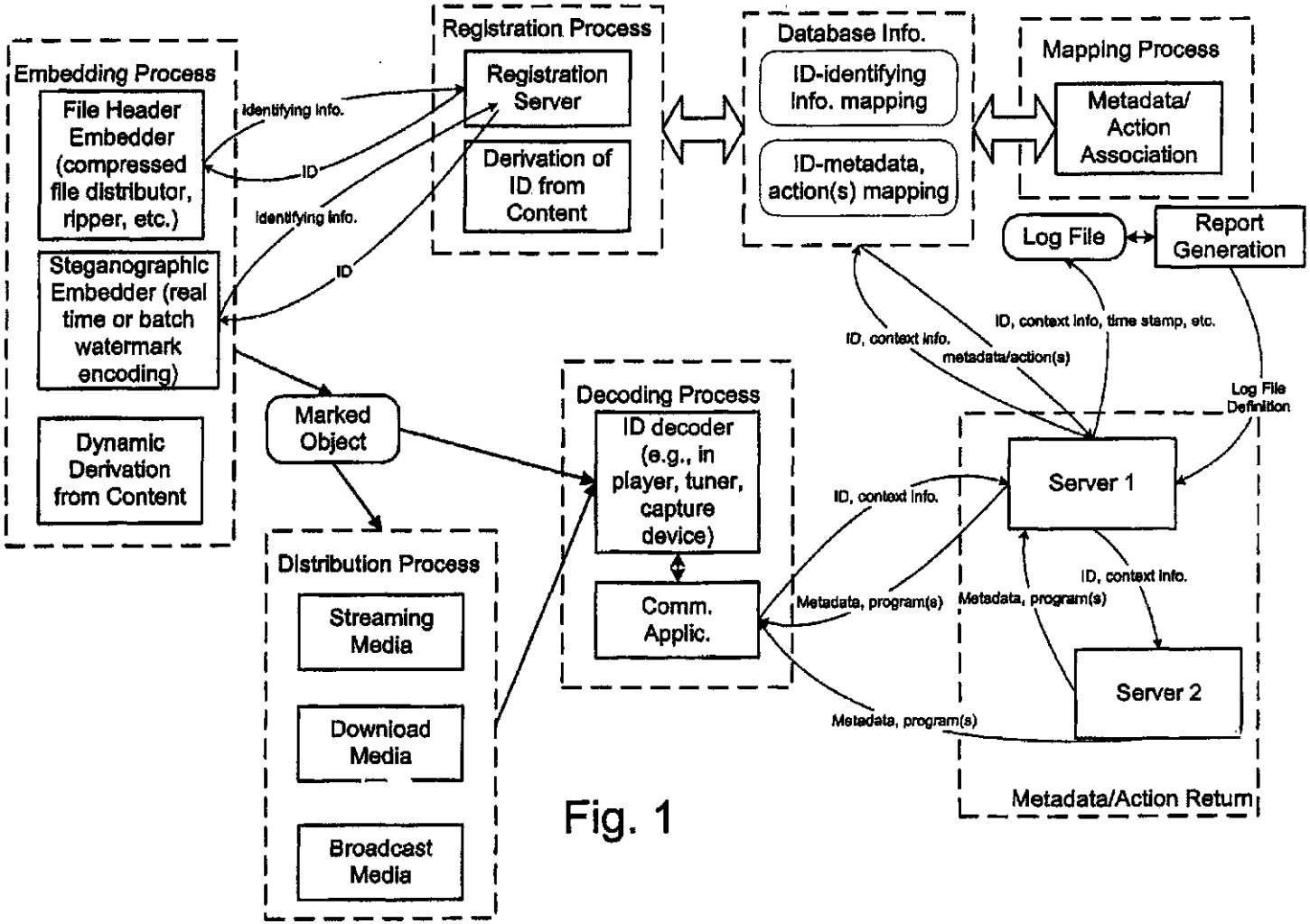

Embedding Process
File Header Embedder (compressed file distributor, ripper, etc.)
Steganographic Embedder (real time or batch watermark encoding)
Dynamic Derivation from Content
Identifying Info.
ID
Identifying Info.
ID
Registration Process
Registration Server
Derivation of ID from Content
Database Info.
ID-identifying Info. mapping
ID-metadata, action(s) mapping
Mapping Process
Metadata/ Action Association
Log File
Report Generation
ID, context info, time stamp, etc.
ID, context info.
metadata/action(s)
Log File Definition
Marked Object
Decoding Process
ID decoder (e.g., in player, tuner, capture device)
Comm. Applic.
ID, context info.
Metadata, program(s)
Server 1
ID, context info.
Metadata, program(s)
Server 2
Metadata, program(s)
Metadata/Action Return
Distribution Process
Streaming Media
Download Media
Broadcast Media

Fig. 1

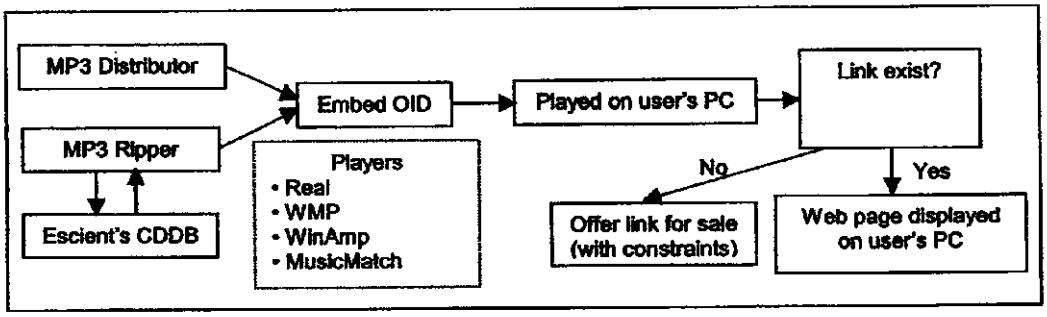

MP3 Distributor
MP3 Ripper
Escient's CDDB
Embed OID
Players
• Real
• WMP
• WinAmp
• MusicMatch
Played on user's PC
Link exist?
No
Yes
Offer link for sale (with constraints)
Web page displayed on user's PC

Fig. 2

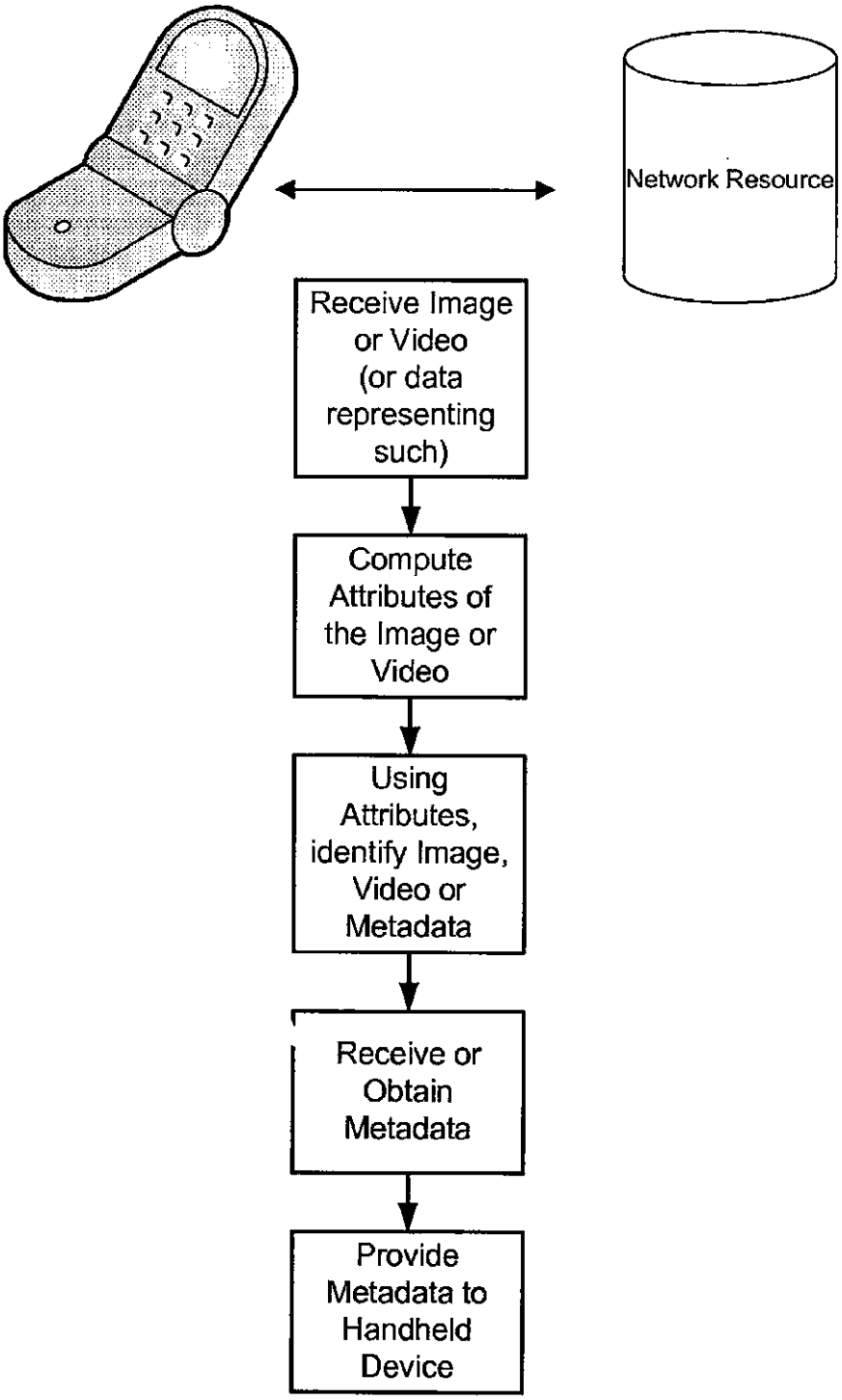

Network Resource
Receive Image or Video (or data representing such)
Compute Attributes of the Image or Video
Using Attributes, identify Image, Video or Metadata
Receive or Obtain Metadata
Provide Metadata to Handheld Device

Fig. 3

DERIVING ATTRIBUTES FROM IMAGES, AUDIO OR VIDEO TO OBTAIN METADATA

RELATED APPLICATION DATA

This patent application is a continuation of U.S. patent application Ser. No. 10/338,032, filed Jan. 6, 2003 (now U.S. Pat. No. 7,349,552), which is a divisional of U.S. patent application Ser. No. 09/563,664, filed May 2, 2000 (now U.S. Pat. No. 6,505,160), which is a continuation in part of U.S. patent application Ser. No. 09/476,686, filed Dec. 30, 1999, which claims priority to U.S. Provisional Application No. 60/134,782, filed May 19, 1999. This application is also a continuation of U.S. patent application Ser. No. 10/338,031, filed Jan. 6, 2003 (now U.S. Pat. No. 7,333,957), which is a divisional of U.S. patent application Ser. No. 09/563,664, filed May 2, 2000 (now U.S. Pat. No. 6,505,160), which is a continuation in part of U.S. patent application Ser. No. 08/746,613 filed Nov. 12, 1996 (now U.S. Pat. No. 6,122,403), which is a continuation in part of U.S. patent application Ser. No. 08/649,419, filed May 16, 1996 (now U.S. Pat. No. 5,862,260), PCT Application PCT/US96/06618, filed May 7, 1996, and U.S. patent application Ser. No. 08/508,083, filed Jul. 27, 1995 (now U.S. Pat. No. 5,841,978).

The subject matter of the present application is related to that disclosed in U.S. Pat. No. 5,862,260, and in co-pending applications Ser. No. 08/746,613, filed Nov. 12, 1996 (allowed); Ser. No. 09/343,104, filed Jun. 29, 1999; 60/164,619, filed Nov. 10, 1999; Ser. No. 09/476,686, filed Dec. 30, 1999; Ser. No. 09/503,881, filed Feb. 14, 2000; Ser. No. 09/525,865, filed Mar. 15, 2000; 60/191,778 filed March 24; and Ser. No. 09/547,664, filed Apr. 12, 2000, which are hereby incorporated by reference.

TECHNICAL FIELD

The invention relates to linking audio and other multimedia data objects with metadata and actions via a communication network, e.g., computer, broadcast, wireless, etc.

BACKGROUND AND SUMMARY

Advances in computer and wireless networking, multimedia coding, and higher bandwidth communication links are creating many new ways to distribute and enjoy multimedia content, such as music and movies. Coding formats for audio like MPEG 1 Layer 3 (MP3) have already caused significant changes in music delivery to consumers. Despite the advances in technology, content distributors and broadcasters still need to address how to effectively promote and sell content.

This disclosure describes systems and processes for linking audio and other multimedia data objects with metadata and actions via a communication network, e.g., computer, broadcast, wireless, etc. Media objects are transformed into active, connected objects via identifiers embedded into them or their containers. These identifiers can be embedded by the owner or distributor of the media object, or automatically created from the media object. In the context of a user's playback experience, a decoding process extracts the identifier from a media object and possibly additional context information and forwards it to a server. The server, in turn, maps

the identifier to an action, such as returning metadata, redirecting the request to one or more other servers, requesting information from another server to identify the media object, etc. If the identifier has no defined action, the server can respond with an option for the user to buy the link and control the resulting action for all objects with the current identifier. The linking process applies to broadcast objects as well as objects transmitted over networks in streaming and compressed file formats.

One embodiment includes a method having: receiving an image or video from a handheld device; computing attributes of the image or video, said act of computing operates on data representing picture elements of the image or video, the attributes are to be used to identify the image or video; using computed attributes of the image or video to identify the image or video, or to identify metadata associated with the image or video in a network resource; retrieving metadata associated with the image or video in the network resource; and providing the metadata to the handheld device.

Another embodiment includes a method having: receiving data corresponding to an image or video from a handheld device, the data representing picture elements of the image or picture elements of the video; computing attributes of the data using a processor, said act of computing controls the processor to operate on the data, the attributes are to be used to identify the image or video; using computed attributes of the image or video to identify the image or video or to identify metadata associated with the image or video in a network resource; obtaining metadata associated with the image or video in the network resource; and providing the metadata to the handheld device.

Yet another embodiment includes a method reciting: obtaining data corresponding to media content from a handheld device, the data representing picture elements of an image or video or representing audible portions of an audio signal; computing attributes of the data using a processor, said act of computing utilizes the processor to operate on the data; using computed attributes of the data to identify the media content or to identify metadata associated with the media content; obtaining metadata associated with the media content; and providing the metadata to the handheld device from a network resource.

Still another embodiment includes a system comprising: a communications channel to obtain data corresponding to an media content from a handheld device, the data representing picture elements of an image or video or representing audible portions of an audio signal; a processor to: i) compute attributes of the data; ii) using at least computed attributes of the media content, identify the media content or identify metadata associated with the media content; iii) obtain metadata associated with the media content in the network resource; and iv) control communication of the metadata to the handheld device from a network resource.

Another embodiment includes a system having: a communications channel to receive information from a handheld device, the information having been derived from data representing picture elements of an image or video, the information to be used at least to identify the image or video, the information having been steganographically hidden in the data representing picture elements of the image or video; a processor to: i) identify the image or video, or to identify metadata associated with the image or video, using the information; ii) obtain metadata associated with the image or video; and iii) control communication of the metadata to the handheld device.

Further features will become apparent with reference to the following detailed description and accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a diagram illustrating examples of media object linking processes and systems.

FIG. 2 is a diagram illustrating media object linking applications.

FIG. 3 is a diagram illustrating providing metadata to a handheld device.

DETAILED DESCRIPTION

Linking Audio and Other Media Objects Via Identifiers

The following sections describe systems and processes for linking audio and other media objects to metadata and actions via an identifier. For the sake of illustration, the disclosure focuses on a specific media type, namely audio signals (e.g., music, sound tracks of audio visual works, voice recordings, etc.). However, these systems, their components and processes apply to other types of media signals as well, including video, still images, graphical models, etc. As described further below, an identifier attached to an audio signal is used to connect that signal with metadata and/or programmatic or device actions. In the context of this document, the terms "media object" and "audio object" refer to an electronic form of a media signal and audio signal, respectively. The linking of media signals applies to objects that are transmitted over wire networks (such as a computer network), wireless networks (such as a wireless telephone network), and broadcast (AM, FM, digital broadcast, etc.).

There are a number of ways to associate an identifier with an audio object. One way to associate the identifier is to insert it in the form of a numeric or alphanumeric code (e.g., binary or M-ary code) in the electronic file in which the audio is stored. Another way to associate the identifier is to embed it as auxiliary data in the audio signal using steganographic methods, such as digital watermarking or other data hiding techniques. Yet another way is to derive the identifier from the audio signal, the table of contents, the file system structure, or its container (e.g., an electronic file or physical package for data like flash memory, Digital Versatile Disk (DVD), minidisk, or compact disk (CD). The physical media may have identifying characteristics, such as a unique identifier or encoded n . . . ta, or other attributes from which an identifier can be derived (e.g., CD disk wobble).

When the identifier is associated with metadata or actions, it transforms the media object into a "linked" object. The identifier travels with the object through distribution, including in some cases, through physical distribution in packaged media and through electronic distribution (broadcast or network communication). The identifier may travel within the same band as the audio object, such as a watermark, or via a separate band, such as a file header or footer or separate broadcast band. A decoding device or programmatic process extracts the identifier from the object and uses it to retrieve related data or actions ("metadata"). In the case of an audio object, like a song, the metadata typically includes the title, artist, lyrics, copyright owner, sound recording owner, information about buying or sampling opportunities and URLs to this type of data as well as web sites and other programs and devices. Linked actions include device or programmatic processes for electronically establishing a license, transferring content (either streaming or download), sending an email, recording marketing data about a transaction, etc. The identifier allows a fan of a particular type of music or artist to get more information about the music and to buy more music.

From the perspective of the artists and record labels, the identifier provides an additional opportunity to promote their music and sell content, concert tickets, etc.

In addition, in some implementations where identifier linking transactions are monitored, it enables the vendors of music to gather data about electronic transactions triggered by the link. For example, users of information may choose to provide information about themselves when they register their decoding device or software with the system. A user ID or other context information may then be recorded when the identifier is extracted and used to trigger a transaction. Many entities involved in the distribution of media signals can benefit from the linking capability. Artists can link their music to information about themselves and provide electronic buying opportunities for music, concert tickets, clothing, etc. Rights holding organizations can use the link to inform users about itself and licensing opportunities. In some cases, the link may also be used to monitor playing and distribution of copies of the music. Record labels can link their music to information about the artist, the label, electronic buying opportunities, etc. Electronic retailers can increase sales by linking users to opportunities to sample and buy additional music (via download or streaming delivery over a wire or wireless network). Conventional brick and mortar retailers can use linking to provide information about the music and to provide buying opportunities. Radio stations and other broadcasters can use the linking capability to bring users to their web sites, creating advertising revenue, to provide electronic buying opportunities for music, concert tickets, clothing items, etc. These and other forms of linked metadata and actions may be implemented in various combinations in different application scenarios.

Depending on the application, the identifier may identify the media object in which it is embedded, or entities, things or actions other than that particular media object. One type of identifier is an object ID that identifies an audio object. This identifier may be a number associated with the object, such as its International Standard Recording Code (ISRC). Another type of identifier is distributor ID that identifies the distributor of the audio object. Another type of identifier is a broadcaster ID that identifiers the broadcaster of the audio object. Of course, more than one identifier may be encoded into an audio object or its container. In the event that an object ID is not encoded with an audio object, but instead a distributor or broadcaster identifier is encoded with the object, other context information, such as the time of play back or distribution, location of distribution, etc. may be used to identify the audio object as part of the linking process. An example is a radio station that marks its broadcasts with a station ID and maintains a playlist database with the air times of each audio object. At decoding time, the station ID is extracted and used along with context information such as the air time of the audio object to look up the audio object or its corresponding metadata and actions. This approach enables the linking system to provide audio object specific metadata or actions even without requiring a unique object identifier in every audio object.

System Implementation

FIG. 1 is a diagram of a system configuration of linked media objects. In this configuration, an identifier links audio objects to metadata via an electronic network, such as the Internet, a wireless network, or a broadcast network. As depicted in FIG. 1, an embedding process may be used to encode an identifier in an audio object or its container. In some cases, an embedding process encodes the identifier in the audio file (e.g., a tag in a file header or footer), in the audio

signal (a digital watermark), or in the physical packaging. The identifier may also be derived as a function of the audio signal or other information in the file or physical packaging (e.g., track information on a CD). In the case of dynamically derived identifiers, an embedding process is not necessary because the identifier can be derived from the content at decoding time.

In some application scenarios, the embedding process interacts with a registration process to get an identifier. The embedding process provides information about the object (e.g., a title and artist name, an ISRC, name of distributor, etc.). In response, the registration process provides an identifier and stores a database record of the association between identifier and the object or other information used in decoding to identify the object, such as its distributor or broadcaster. The registration process may be used to assign an identifier to an audio object and to distributors or broadcasters of audio objects. The embedding and registration processes may occur before the audio object is distributed to consumers, or sometime thereafter, such as when a user transfers (e.g., "rips") an a media object from one format to another (e.g., a packaged format to an electronic file format such as a compressed file format).

Once registered, an interactive or automated mapping process associates the identifier with data or actions. The registration process creates a database of identifiers and associates the identifiers with corresponding media objects, distributors, broadcasters, etc. The mapping process associates the identifiers with corresponding metadata or actions.

Once associated with an audio object and metadata, the identifier transforms the audio object into a linked object. The identifier remains with the object through distribution, although some embedding processes are more robust than others to intentional or unintentional distortion/removal of the identifier. There a variety of different distribution scenarios. Some examples depicted in FIG. 1 include transferring an audio object over a computer network, streaming the object over a computer network, or broadcasting it (e.g., AM/FM broadcasting, digital broadcasting, broadcasting over wireless carriers, etc.). Whatever the distribution process, a user ultimately receives the linked object in a player, tuner, or capture device.

To activate the linked object, a decoding process extracts the identifier and uses it to access associated data or actions. The decoding process may be implemented as a separate program or device, or integrated into a player, tuner, or some other capture device, such as listening devices that convert ambient audio waves to an electronic signal and then extract the identifier from the signal.

In the configuration shown in FIG. 1, the decoding process forwards the extracted identifier to a communication application, which in turn, forwards it in a message to a server. The decoding process or the communication application may add additional context information to the message sent to the to a server. The context information may relate to the user, the user's device, the attributes of the session (time of playback, format of playback, type of distribution (e.g., broadcast or transmitted audio file), etc.) Based on identifier and optional context information, the server determines an associated action to perform, such as re-directing an identifier or context data to another server, returning metadata (including programs, content, etc.), downloading content, logging a transaction record. To find the associated action or actions, the server maps the identifier to actions based on the information established in the mapping process. The server may: 1) look up the data and actions in a local database stored in its memory subsystem; 2) route the identifier to one or more other servers via the network, which in turn look up related actions and data associated with the identifier; or 3) perform some combination of actions 1 and 2.

In the first case, server 1 returns data or actions associated with the identifier. The server may look up related data based on the identifier alone, or based on the identifier and other context information. Context information may be information provided by the user, by the user's computer or device, or by some other process or device. In the second case, the server looks up one or more addresses associated with the identifier and forwards the identifier and/or possibly other context data to secondary servers at these addresses via conventional networking protocols. Again, this context data may include data from the user, the user's computer, some other device or database. For example, server 1 might query a remote database for instructions about how to process an identifier. These instruction may specify data to return to the communication application or to forward to another server, which in turn, looks up associated data and returns it to the communication application. A server may return data that an audio player displays to the user or uses to control rendering of the content. For example, the server can tell the player that the object contains inappropriate content for children. The player or user can make decisions about whether or how to play the material based on this information.

Both the server and the player can adopt a set of rules. The server rules may be used to control what the server returns in response to an identifier and context data. The player rules may be used to control what the player displays to the user or how it renders the content based on data returned from a server.

Either the first server, or a server one or more levels of indirection from the identifier may return data and programmatic actions to a player via the communication application. Each server in these levels of indirection receives a database key, such as an identifier or context information, from the previous server, and uses it to look up corresponding actions. These actions may include returning data or programs to the communication application or to previous servers in the routing path of the message from the communication application. Also, the servers may route requests for information or actions to other servers. The server or servers may return data or perform actions in response to the identifier (or other context data) that do not directly impact the decoding process, or the device in which it o[illegible]s.

The system depicted in FIG. 1 allows several different interested parties to establish services linked via the identifier. For example, server 1 can be configured to provide generic promotional and/or licensing information associated with an identifier. If the content owner, distributor, retailer, artist or other related party wishes to provide information or services for a connected object, then server 1 may also route the identifier for that object, and possibly context information, the address of the communication application, and instructions, to servers maintained by these entities. These servers, in turn, provide promotional, sales, or licensing information, and electronic buying or licensing opportunities specific to that entity back to the consumer over the network via the communication application.

In the context of a network configuration, Internet protocols may be used to return data to the communication application or to the device or system in which it operates. The communication application may be implemented in a web browser, such as Internet Explorer or Netscape Navigator. Examples of ways of exchanging information between a client player and a server include returning a web page with metadata and program scripts designed to run on the end

user's system. The metadata itself may include active links, such as URLs to other network resources, such as a web site or some other network service. The path of the identifier from the decoding process, and the return path from a server to the communication application may include one or more hops through a wire or wireless connection using standard wire and wireless communication protocols like TCP/IP, HTTP, XML, WAP, Bluetooth, etc. In addition, data returned to the user may be routed through one or more servers that may forward the data, and in some cases, augment the data or modify it in some fashion.

FIG. **2** is a diagram illustrating applications of the system depicted in FIG. **1**. In the application scenarios depicted in FIG. **2**, an embedding process encodes an object identifier (OID) into an audio file, such as an ID3 tag in the header of an MP3 file or audio frame headers in the MP3 file. FIG. **2** shows two embedding scenarios. The first is an MP3 distributor that embeds OIDs in MP3 files before transmitting them over a network, such as the Internet, typically via a web site interface. The second is a file ripping process where a programmed computer or other device extracts an audio object from packaged media such as a CD and converts it into a coded file format like MP3. In the latter case, the ripping process may extract metadata from the CD, such as the table of contents, and use this metadata as a key to a database (CDDB) to get information about the songs on the CD, such as title, artists, etc. The table of contents or other metadata from a package medium, such as optical or magnetic storage or flash memory, may be hashed into an index to a database entry that stores information about the media signal stored on the medium. The ripping process uses the information returned from the database to identify the audio objects on the packaged media so that they can be associated with an OID. This is an example of identifying information used to associate an OID with an audio object. As part of the coding process, the ripping process inserts the OID in the file header of the MP3 file.

Later, when a user opens or plays the marked MP3 in a player, such as a software player like the real player, Liquid Audio player, Windows Media Player (WMP), WinAmp, MusicMatch, etc., a plug-in software module in the player extracts the OID and forwards it to a server via an Internet connection. The plug-in may establish its own Internet connection, or pass the OID to an Internet Browser, which in turn, establishes a connection (if one is not already present) with the server. As an intermediate step, the plug-in may display a window with user options, such as "learn more about the song", "play the song", or both. The user can then choose to get more information by actuating the first or third options in the user interface window, which cause the plug-in to forward the OID to the server.

The server then returns a web page associated with the OID, or re-directs the OID to another server (e.g., one maintained by the content distributor or owner), which in turn, returns a web page of information about the object and links to related actions (e.g., a link to a licensing server, a link to a server for buying and downloading related music etc.). The licensing server may be programmed to download software players and new music offerings compatible with those players. For instance, the licensing server may provide software for decrypting, decoding, and playing electronically distributed music according to usage rules packaged with the electronically distributed music. In this application scenario, the linking of the MP3 file enables the content owner to market music and products that promote the sale of audio objects in other formats, included formats protected with encryption, watermark copy managements schemes, etc.

In the event that a media object is not linked, the decoding and server processes can be programmed to enable the user to purchase a link for the object. For example in one scenario, the player plug-in displays a graphic for a link information indicating that the link is available after determining that an OID is not in the file. If the user clicks on the graphic, the plug-in displays more information about the procedure for purchasing or renting a link. This information may be provided in conjunction with querying the server and displaying information returned from the server, or alternatively, providing pre-programmed information incorporated into the plug-in. If the user is interested in purchasing the link, he or she can then enter input (e.g., click on a button such as "Get Link") that initiates the process of registering an OID with the object and associating metadata or actions with the OID. The process of registering the OID and associating the OID with metadata or actions may be performed as described in this document. This scenario provides yet another mechanism for transforming content into connected content.

There are many possible variations to the applications scenarios illustrated in FIG. **2**. During the file ripping process (or some other embedding process), the embedder may generate a unique ID from the metadata read from the packaged media on which the media object resides. One example of such an ID is the number derived from CD metadata currently used to index information in the CDDB database. This ID may then be embedded in the audio object or its file header/footer. During OID registration, the registration process may inform the embedding process that the OID (and thus, the object for which it was derived) has not been associated with metadata or actions. In this case, the user may be given an opportunity to purchase the link, either at the time of ripping, or in the future, wherever the object travels. In the latter case, the OID in the object is associated with an option to buy the link and customize the data and/or actions associated with that link. Rather than link to promotional information, the OID gives users an option to buy or rent the link and provides them with an opportunity to customize it (e.g., linking it to a custom web site). Once customized, other users that open or play the file will then be able to link to the customized information or actions.

To assert control over the type of customization that users may perform, the registration and mapping processes can place constraints on the types of metadata and actions that users can link to a media object.

In the multimedia content industry, there are typically many rights holders and entities involved in the distribution process. This may present a conflict when linking a media object to one entity. One way to address this problem is have an object link to many different entities. For example, the server could map an OID to many entities and return links to retailers, distributors, record labels and artists. Another way to address it is to encode additional information about the distributor in the OID. For example, the OID includes fields that identify the object and its distributor. If a user activates the link to purchase products, including media objects, then the distributor name is logged with the purchase and that distributor is credited with royalties associated with the transaction. The distributor field may also be used as a key to look up the appropriate action for the OID, such as re-directing the OID to the web server of the entity associated with that OID. In this approach, even if the OID directs a user to a record label's website, the distributor field can be used to credit the distributor with a royalty for the linking transaction.

The entity responsible for maintaining a web site linked via on identifier can make deals with online resources for providing data about a media object such as lyrics, song titles, radio

station play lists. The website may link to this information, access it via a database manager, etc.

File Identifiers

One form of identifier is an identifier that is inserted in an audio object file, but in a distinct field from the audio signal itself. Some examples are file headers and footers. This file identifier may be assigned before or after distribution of the audio object to consumers. In addition, it may be derived from the audio signal or other information in the file. For example, an identifier generator may derive a unique or sufficiently unique identifier from a portion of a music signal. A variety of methods for generating a unique numbers based on a unique collection of numbers may be used.

The process of embedding a file identifier may be done at the time of encoding or transcoding a file. For example, the file identifier may be inserted during a ripping process, such as when a device or programmatic process converts a song from a format stored on packaged media, like a CD or DVD, to an electronic, and compressed form, such as MP3 or some other audio codec. As another example, the file identifier may be inserted when a device or programmatic process transcodes an electronic music file from one codec format to another. Yet another example is where a file is taken from a digital or analog uncompressed format, and placed in another format for distribution.

Identifiers Embedded in Audio Signal

Another way to associate an identifier with an audio signal is to embed the identifier in the audio signal using steganographic methods, such as digital watermarking or other data hiding techniques. Many of such techniques have been developed and are described in published articles and patents. Watermarking methods are described in U.S. patent application Ser. No. 09/503,881. Other examples of methods for encoding and decoding auxiliary signals into audio signals include U.S. Pat. Nos. 5,862,260, 5,940,135 and 5,945,932. For more information on steganographic applications, see the patent applications incorporated by reference.

The steganographic embedding method may be performed in a batch process. Consider a distributor of electronic music via the Internet or some other network, or a broadcaster of music such as a radio station. In each case, the distributor and broadcaster have a collection of audio objects. The embedding process may operate on a collection of objects in a batch process by retrieving an electronic version, encoding an identifier obtained from the registration process, and returning the marked version for later distribution or broadcasting. In some cases, it is desirable to do watermark embedding in an iterative process in a studio environment to encode the watermark with an intensity that achieves desired perceptibility and robustness requirements.

The steganographic embedding method may also be performed at the time of transmission of an electronic file or broadcast of the audio object. In the case of distribution via a network such as the Internet (e.g., streaming or file download), real time embedding enables the embedding process to also embed context information that is specific to the consumer (or the consumer's computer) that has electronically ordered the object. For example, when the user requests a file in a streaming or a compressed file format via the Internet using her browser, the distributor's server can request information (perhaps voluntary) about the user to be associated with the transmitted object. Later, the decoding process or the servers that map the identifier to actions or metadata can use this information to determine the types of information to provide or responsive action to perform.

In the case of broadcasting, real time embedding enables the identifier to be steganographically embedded throughout an electronic version of the audio signal just before, or as part of the broadcasting process.

An object or distributor ID (as well as other identifiers or context information) can be embedded in the payload of a watermark that is also used for copy control. Portion of the watermark can be used to control whether the object can be played, transferred, recorded, etc., while another part can be used to carry identifiers and other metadata for linking functions described in this document. Alternatively, entirely separate watermark encoding and decoding methods may be used for copy control and linking functions.

A watermarking process may be used to encode different watermarks in the various channels of an audio signal. Message information may be embedded in one or more channels, while synchronization or orientation signals used to detect and decode the message information may be encoded in other channels. Also, different messages (e.g., different identifiers) may be encoded in different channels. At decoding time, the different identifiers can trigger different actions or link to different data.

In broadcasting applications, an identifier may be encoded along with the broadcast of the associated media signal by modulating a subcarrier of the main carrier frequency used to transmit the media signal. The subcarrier conveys auxiliary data such as the identifier, while the main carrier conveys the associated media signal To reduce audibility of the auxiliary data (e.g., the identifier(s)) encoded in the sub-carrier, the data can be randomized by applying it to a pseudorandom or random number by some function that may be inverted in the decoding process, e.g., multiplication or exclusive OR functions. One example of sub-carrier encoding and decoding is Active HSIDS 97 developed by Seiko Corporation.

Identifiers in Digital Radio Broadcasts

Some forms of digital radio broadcasts support transmission of metadata along with media signals. This metadata can also be used to carry one or more identifiers that are mapped to metadata or actions. The metadata can be encoded at the time of broadcast or prior to broadcasting. Decoding of the identifier may be performed at the digital receiver. In particular, the digital receiver receives the broadcast data, extracts the identifier, and either automatically, or at the user's direction, forwards the identifier to a server to look up the associated metadata or action.

Dynamic Identifier Extraction from Audio Content or Related Data

As noted above, another way to associate an identifier with a corresponding audio signal is to derive the identifier from the signal. This approach has the advantage that the embedding process is unnecessary. Instead, the decoding process can generate the identifier from the audio object. In this case, the decoder computes a fingerprint of the audio signal based on a specified fingerprinting algorithm, The fingerprint is a number derived from a digital audio signal that serves as a statistically unique identifier of that signal, meaning that there is a high probability that the fingerprint was derived from the audio signal in question. One component of fingerprint algorithm is a hash algorithm. The hash algorithm may be applied to a selected portion of a music file (e.g., the first 10 seconds) to create a fingerprint. It may be applied to discrete samples in this portion, or to attributes that are less sensitive to typical audio processing. Examples of less sensitive attributes include most significant bits of audio samples or a low pass filtered version of the portion. Examples of hashing algorithms include MD5, MD2, SHA, SHA1.

As an aside, fingerprinting may also be used to determine whether an audio signal has been watermarked. The fingerprinting application can evaluate a fingerprint for a received object and compare it with one for a watermarked object (or unmarked object) to determine whether the object is likely to be watermarked. Certain fingerprints can be associated with certain types of watermark methods. Using the fingerprint, a decoding device can select an appropriate watermark decoding system for the object.

While specifically discussed in the context of audio objects, the fingerprinting process applies to other types of multimedia content as well, including still images, video, graphics models, etc. For still images and video, the identifier can be derived dynamically from a compressed or uncompressed version of the image or video signal. The fingerprinting process may be tuned to generate a specific identifier based on the type of file format. For example, the process extracts the file format from the file (e.g., from a header or footer), then uses a fingerprinting process tailored for that type of file (e.g., a hash of a compressed image or video frame). The dynamic identifier computed by this process may be associated with metadata and/or actions using the processes and systems described in this document.

Registration Process

One way to implement the registration process is to build client and server application programs that communicate over a computer network using standard network communication protocols. The client may be implemented as a software program that provides identifying information about an audio object. It can obtain the information by prompting the user for the identifying information, or from extracting it from the audio object or its container. The server may be implemented as a database management program that manages identifiers and corresponding audio objects. When queried to provide an identifier for particular identifying information, the program checks whether it has already assigned an identifier to an object based on the identifying information. If so, it returns that identifier that has already been assigned. If not, it assigns a new identifier number, creates a new entry in the database for that number and its associated identifying information.

The type of identifier used to link audio objects varies with the application. As such, the registration process may vary as well. One type of identifier is a unique identifier for [illegible]udio object. Another type of identifier is one that identifies some attribute of the audio object, but does not uniquely identify it, such as a distributor or broadcaster identifier. This type of identifier requires additional context information to uniquely identify the audio object at the time of linking it to actions or metadata. For these types of identifiers, the registration process provides information identifying the attribute of the audio object, such as its distributor or broadcaster. In response, the server provides an identifier that may be embedded in several audio objects that share that attribute.

One example is a broadcaster ID, such as a radio station ID. Audio broadcast by the radio station is embedded with this radio station ID. To identify the object, context information such as the play time captured at the tuner is used along with the radio station ID extracted from the received audio signal to identify the audio object. The decoding process forwards this information to a server. Using the radio station ID and context information, the server maps the ID to an appropriate action. This may include querying a radio station's playlist database for an object identifier based on the station ID and context information. The server can then map the object identifier to an action or metadata based on the object ID returned from the playlist database. Other scenarios are possible. For example, the server could forward the station ID, context data and decoder address to a radio station server, which in turn, looks up the appropriate action or metadata (e.g., web page) and sends it to the device that decoded the station ID.

Broadcast content can also be associated with object identifiers. One way to implement the identifier assignment process is to allocate a unique set of identifiers with each broadcaster/distributor. Those broadcasters or distributors are then free to assign the identifiers to media objects as they wish. Once they complete the identifier assignment process, they may then associate the identifiers with the metadata or actions in a mapping process.

Embedding Process

The embedding process may be integrated into a software program along with the client of the registration process described in the previous section. This integration of registration and embedding functions is particularly suited to a batch embedder, where processing time required to request an identifier is less of a concern.

In real time embedding, the identifier or identifiers are preferably available for associated audio objects before embedding begins. For example, the identifiers can be maintained in a local database on the embedding computer or device and indexed by object title. Distributor and broadcast identifiers are more straightforward because they may be applied to several different audio objects.

The embedding process may also be implemented in an embedding clearinghouse system. The embedding clearinghouse is a computer or other electronic system that analyzes media objects and embeds one or more links in the media objects. The clearinghouse may be implemented in a server on a network, such as the Internet and operate on content in a "push," "pull," or some combination of push and pull models. In the push model, users and other systems send media objects to the embedding clearinghouse for analysis and embedding. The pull model, the clearinghouse has the capability to search for and gather media objects for embedding and analysis. One example of this pull model is an Internet search process called a spider that crawls the Internet, searching for media objects to analyze and embed with one or more identifying links.

The embedding clearinghouse analyzes a media object (perhaps based on out of band data like a file header or footer) and inserts an identifier. This identifier may link [illegible] metadata and actions, such as re-direction to a web site offering products, services, and information related to the content. The embedding clearinghouse may incorporate search engine technology to execute a key word search based on information from the media object and then associate the media object with a series of related URLs returned from the Internet search. The process may be automatic, or with some user input to select which sub-set of links should be inserted.

The embedding clearinghouse may also offer an identifier embedding services for those wanting to link their media objects with metadata, actions, etc. In this application scenario, the embedding clearinghouse may be implemented as an Internet server that is accessible via a web page using conventional network communication and web protocols. To access the server, users visit a web page using an Internet browser. In exchange for a fee, which may be tendered electronically over the Internet from the user's computer to the server, the server provides an embedding service to embed an identifier into a media object uploaded from the user via the user's computer and Internet connection. The user can select the information to associate with a media object, such as generic identifying information (e.g., title, author, owner),

generic licensing information, or special information or actions. The generic information is hosted by the provider of the embedding clearinghouse server, while the special purpose information and actions are accessed through re-direction. In particular, the provider of the clearinghouse server links the embedded identifier to an address or set of addresses of servers that provide the special information or actions. Then at decoding time, the decoding process sends the identifier to the provider's server, which in turn, redirects the identifier to a secondary server or servers that provide special purpose information or actions (e.g., redirect to a web page of the content owner, download related content, provide electronic licensing services, etc.).

Decoding the ID and Embedded Context Data

The implementation details of the decoding process depend on how the identifier is encoded into an audio object or its container. In the case where the identifier is encoded in a file header or footer, the decoder may be a software program or digital hardware that parses the header/footer and forwards it to the communication application. One way to implement this type of decoder is to integrate it into a media player as a plug in program. Examples of media players include Windows Media Player from Microsoft, Liquid Audio player from Liquid Audio, Winamp, Real Player from Real Networks. Preferably, the plug-in gives the user visual feedback that the identifier has been detected and displays a window with options to access more information or actions available via the link. For example, the user can be presented with a user interfaces prompting the user to click for more information or buying opportunities. If the user selects these options, the plug-in forwards the user selections and identifier to the communication application, which forwards them to the server (e.g., server **1**, FIG. **1**).

In the case where the identifier is steganographically encoded in the audio object, a corresponding decoder extracts the identifier. This type of decoder may be implemented as a plug in to a software player as described in the previous paragraph. It may also be implemented in a tuner for broadcast content, or in a listening device that captures audio from the ambient environment.

In the case where the identifier is derived from the content or container metadata, the decoder captures the pertinent portion of the audio object, and generates the identifier as described above. This type of decoder can be implemented in a software or hardware player, a tuner, etc.

The decoder may collect identifiers in response to a user request while objects containing these identifiers are being played. For example, when the user is playing music, he may like a song and want to buy it or get more information. This feature may be implemented by building an interface that has a button or voice recognition that enables the user to request information or a buy/license opportunity. Once captured, identifiers can be forwarded along with user instructions to the appropriate server.

However, one particularly useful feature is to enable the user to fetch information and make orders from music as the music is playing. The system described previously supports this feature because the decoding process can forward the identifier or identifiers, embedded context information, or additional context information (user information, play time, broadcast type, file type, player type, operating system type) to the communication application as the music is playing. The user can trigger the linking action by pressing a "fetch" button, or saying fetch to a voice activated input device that causes the decoding device to package a message and invoke the communication application (e.g., Internet browser). In turn, the communication application forwards the message to a server that parses the message and determines the associated action.

The activation of the "fetch it" feature may be made on a handheld device that communicates with a decoding device in a tuner via a wireless connection. For example, a user may press a button on a remote control device, like a key chain, which sends a wireless signal to a receiver in the tuner. The receiver invokes the decoding process. The tuner may also send metadata from the server to the remote control device for display using a similar wireless connection. Infrared or RF transceivers, for example, may be used to communicate the data back and forth.

The decoding device may also provide continuous decoding of identifiers. When the user requests a "fetch," the identifier and context information for the current song may be forwarded to the server. Also, the decoding device may automatically fetch generic information such as song title and artist so that this information is immediately available to the user.

Another possible implementation is to temporarily buffer identifiers extracted from some predetermined number of the most recent songs, titles, etc. These identifiers can be stored along with other metadata, such as a time stamp, to inform the user when they were captured. The user can then select one or more of the items to send to the server for more information or related actions.

These features may be implemented in one or more devices. While the example above discusses a remote control device and a separate tuner with a decoder, these functions may be integrated into a single device, such as a car stereo, phone handset, personal digital assistant, and a variety of other types of players or tuners.

The identifier enables dynamic linking. Dynamic linking enables the identifier encoded with a media object to remain fixed, while the metadata or actions associated with that identifier can be changed. To change the associated metadata, the mapping process edits the identifier database to associate new metadata or actions with an identifier. The mapping process can be automated to change metadata or actions associated with an identifier at periodic intervals or in response to system events. In addition, a user may change the associated metadata or actions interactively at any time. To facilitate access to the database, a web based interface can be added to the database.

Dynamically linked data returned from a server to a player environment can be displayed to the user in a variety of ways. One way is to display it in a web page or user interface window of a player. The data can be animated by scrolling it across the visual display. The data can also be displayed in the form of HTML links, which, when activated, cause the download of other data or initiate actions, such as playing streaming content from a server.

Server Types

As discussed elsewhere, the servers used to link identifiers to actions may be programmed to provide a variety of actions including:

- returning data and HTML links (e.g., in the form of an HTML document, scripts, etc.)
- downloading media signals in streaming or file format
- performing an electronic transaction (selling products like CDs, DVDs, concert tickets, etc. via computer transaction using credit cards, digital money, etc.)

establishing a license to use a linked media object
re-directing to another server
performing database look up operations for related information, links, actions
performing database look up to uniquely identify a media object based on distributor/broadcaster ID and other context information
creating a transaction log

This is by no means in exhaustive list. Another type of server action is to initiate a process of searching a database, a collection of databases or the Internet for additional information related to a linked media object. This type of search service may be performed continuously and the results associated with the identifier. Then, in response to a request from a decoding process, the server can return a digest of the results with links to web pages for additional information.

Communication Application

The implementation details of the communication application are highly dependent on the type of communication link and protocols used to connect the decoding process to a server. Above, an Internet browser is provided as an example. A browser may be implemented in conventional PCs, handheld devices, wireless phones, stereo systems, set top boxes, etc. However, the communication application need not be based on computer network protocols. For wireless devices, where the marked content is played on wireless carrier frequencies, the communication application can employ wireless communication technology to forward identifiers and context information to servers that map this information to actions or metadata and return it via a wireless carrier frequency to user's handset.

Tracking Transactions and Report Generation

As depicted in FIG. 1 and described above, the servers for mapping identifiers to actions may be programmed to dispense a transaction log into a log file. A report generation process can then enable users to define and request queries of data from the log file based on a particular identifier, a particular type of context information (time frame, geographic location, user demographics, etc.), a particular action, etc.

Capture Devices

As noted above, the decoding process may be implemented in a variety of devices or software that process media objc
These devices an ...ware include programmable devices such as personal computers, personal digital assistants, telephone handsets, set-top boxes, personal stereos, hi-fi components, tuners, receivers, televisions, etc. as well as hardwired devices that may be incorporated into these systems and devices.

In some contexts, it is useful to implement a recording function. This is particularly true in devices that receive a broadcast or stream of media content and need to capture at least a portion of it to decode an identifier. Examples of these devices are radio receivers, and wireless telephone handsets. The record function may be automatic or user activated. In the latter case, the user actuates an input device to control the record process and optionally the record duration. For example, the user may hear a song that she likes and press record. The device, in turn, records at least a part of the object that is currently being received (an audio, visual or audio visual signal). The user can then decide contemporaneously or at a later time to execute the identifier decoding process on the recorded signal. The recording function can be designed to execute for a pre-determined or user specified duration.

In the case of radio and television tuners/receivers, the record function can be used to capture a media signal as it is received. In the case of a telephone handset, the record function can be used for a variety of functions, such as recording part of a telephone conversation, recording speech or other ambient audio through a microphone, or recording a media signal received by the handset via a wireless communication channel. The recordings can be compressed and stored in local memory on the device. In addition, they may be annotated with metadata about the media signal, such as a time stamp to show time of capture, a location stamp to show location of capture, metadata extracted from the object (in band or out of band data), etc. The location stamp may be provided by a global positioning device. Some wireless phone systems are capable of computing location of a telephone handset via triangulation. This location data may be used to provide geographic location coordinates or the name of nearby landmark, city name, etc.

The metadata may be displayed on a display device to help the user remember the context of a particular recording. In addition, it may be provided as context information along with an identifier to a server that links the identifier and context information to metadata or actions.

Transmarking

In some applications, it may be useful to convert auxiliary information embedded in a media signal from one format to another. This converting process is referred to as transmarking. Transmarking may include converting an out of band identifier like a tag in a header/footer to a watermark or vice versa. It may also involve converting a message in one watermark format to another. The process involves a decoding operating on an input media object, and an encoding of the decoded information into the media object. It may also involve a process for removing the mark originally in the input object to avoid interference with the newly inserted mark.

There are a variety of reasons to perform transmarking. One is to make the embedded information more robust to the types of processing that the media object is likely to encounter, such as converting from one watermark used in packaged media to another watermark used in compressed, and electronically distributed media, or a watermark used in radio or wireless phone broadcast transmission applications.

This type of transmarking process may be performed at various stages of a media object's distribution pat... s suggest previously, an identifier in a watermark or file header/footer may be encoded at the time of packaging the content for distribution, either in an electronic distribution format or a physical packaged medium, such as an optical disk or magnetic memory device. At some point, the media signal may be converted from one format to another. This format conversion stage is an opportunity to perform transmarking that is tailored for the new format in terms of robustness and perceptibility concerns. The new format may be a broadcast format such as digital radio broadcast, or AM or FM radio broadcast. In this case, the identifier may be transmarked into a watermark or other metadata format that is robust for broadcast applications. The new format may be a compressed file format (e.g., ripping from an optical disk to an MP3 format). In this case, the identifier may be transmarked into a file header/footer or watermark format that is robust and compatible with the compressed file format.

The transmarking process may leave an existing embedded identifier in tact and layer an additional identifier into the media object. This may include encoding a new watermark that does not interfere with an existing watermark (e.g., insert the new watermark in unmarked portions of the media object

or in a non-interfering transform domain). It may also include adding additional or new identifier tags to headers or footers in the file format.

Amplifying an Embedded Identifier

Rather than converting embedded data to another format, an amplifying process may be used to renew an identifier that has become weakened or separated due to processing of the media object in which it is embedded. In this case, an decoder and encoder pair may be used to determine the current identifier and re-encode it. Of course, the encoder can also choose to embed a new or additional identifiers as well.

If the previous identifier is lost, the encoder can query an identifier database established in the registration process, passing identifying information about the media object. The database uses the identifying information to find an associated identifier and returns it to the encoder for embedding in the media object.

CONCLUDING REMARKS

Having described and illustrated the principles of the technology with reference to specific implementations, it will be recognized that the technology can be implemented in many other, different, forms. To provide a comprehensive disclosure without unduly lengthening the specification, applicants incorporate by reference the patents and patent applications referenced above. These patents and patent applications provide additional implementation details. They describe ways to implement processes and components of the systems described above. Processes and components described in these applications may be used in various combinations, and in some cases, interchangeably with processes and components described above.

The particular combinations of elements and features in the above-detailed embodiments are exemplary only; the interchanging and substitution of these teachings with other teachings in this and the incorporated-by-reference patents/applications are also contemplated.

We claim:

1. A method comprising:
   - receiving an image or video from a process in a handheld device;
   - computing attributes of the image or video, said act of computing operates on data representing picture elements of the image or video;
   - using said computed attributes of the image or video to identify the image or video or to identify metadata associated with the image or video in a network resource;
   - retrieving metadata associated with the image or video in the network resource; and
   - providing the metadata for output or display.

2. The method of claim 1, wherein the metadata comprises metadata selected from a group of metadata comprising: audio, video, an image and a URL.

3. The method of claim 1, wherein the metadata comprises at least one action.

4. The method of claim 1, wherein the metadata comprises purchasing information.

5. The method of claim 1, wherein said computing attributes comprises deriving an identifier from the data representing picture elements of the image or video itself.

6. The method of claim 1 wherein the attributes of the image or video is steganographically hidden in data representing picture elements of the image or video.

7. A computer readable medium including instructions stored thereon to execute the method of claim 1.

8. The method of claim 1 wherein said act of computing attributes of the image or video is carried out on an electronic processor configured to execute said act of computing attributes of the image or video.

9. The method of claim 1 wherein the image or video is received from ripping the image or video from a CD or DVD.

10. A programmed computing device storing instructions in memory, said instructions are executable by said programmed computing device to perform the acts of claim 1.

11. A computer readable media comprising instructions stored thereon to cause a multi-purpose electronic processor to perform the acts of claim 1.

12. The method of claim 1 in which a recognition unit performs said act of computing attributes of the image or video.

13. The method of claim 12 in which the recognition unit communicates with video instrumentation.

14. The method of claim 13 in which the video instrumentation comprises a television.

15. The method of claim 1 in which the handheld device comprises a recognition unit for performing said act of computing attributes of the image or video.

16. The method of claim 15 in which the recognition unit communicates with video instrumentation.

17. The method of claim 16 in which the video instrumentation comprises a television.

18. The method of claim 1 in which the image or video is captured with a camera.

19. A method comprising:
   - receiving information from a handheld device, the information having been derived from data representing picture elements of an image or video, the information to be used at least to identify the image or video, the information having been steganographically hidden in the data representing picture elements of the image or video;
   - using the information to identify the image or video, or to identify metadata associated with the image or video, with reference to at least one network resource;
   - retrieving metadata associated with the image or video from the network resource; and
   - providing the metadata to the handheld device.

20. The method of claim 19, wherein the metadata comprises metadata selected from a group of metadata comprising: audio, video, an image and a URL.

21. The method of claim 19, wherein the metadata comprises at least one action.

22. The method of claim 19, wherein the metadata comprises purchasing information.

23. A computer readable medium including instructions stored thereon to execute the method of claim 19.

24. A programmed computing device storing instructions in memory, said instructions are executable by said programmed computing device to perform the acts of claim 19.

25. A computer readable media comprising instructions stored thereon to cause a multi-purpose electronic processor to perform the acts of claim 19.

26. A method comprising:
   - obtaining data corresponding to media content from a handheld device, the data representing picture elements of video or representing audible portions of an audio signal;
   - computing attributes of the data using a configured multi-purpose processor, said act of computing utilizes the configured multi-purpose processor to operate on the data;

using computed attributes of the data to identify the media content or to identify metadata associated with the media content;

obtaining metadata associated with the media content; and

providing the metadata to the handheld device from a network resource.

27. The method of claim **26** wherein the multi-purpose processor is located remotely with respect to the handheld device.

28. The method of claim **26** wherein the attributes are steganographically hidden in the data as obtained.

29. A computer readable medium including instructions stored thereon to execute the method of claim **26**.

30. A programmed computing device storing instructions in memory, said instructions are executable by said programmed computing device to perform the acts of claim **26**.

31. A computer readable media comprising instructions stored thereon to cause a multi-purpose processor to perform the acts of claim **26**.

32. A system comprising:

a communications channel to obtain data corresponding to media content from a handheld device, the data representing picture elements of video or representing audible portions of an audio signal;

an electronic processor configured to: i) compute attributes of the data; ii) using at least computed attributes of the media content, identify the media content or identify metadata associated with the media content; iii) obtain metadata associated with the media content in the network resource; and iv) control communication of the metadata to the handheld device from a network resource.

33. The system of claim **32** wherein the attributes are steganographically hidden in the data as obtained.

34. The system of claim **32** wherein said electronic processor is operating to perform at least one function recited therein.

35. A system comprising:

a communications channel to receive information from a handheld device, the information having been derived from data representing picture elements of an image or video, the information to be used at least to identify the image or video, the information having been steganographically hidden in the data representing picture elements of the image or video;

a multi-purpose processor configured to: i) identify the image or video or identify metadata associated with the image or video, using the information; ii) obtain metadata associated with the image or video; and iii) control communication of the metadata to the handheld device.

36. The system of claim **35** wherein said multi-purpose processor is operating to perform at least one function recited therein.

37. One or more computer readable media comprising instructions stored thereon that when executed by one or more processors perform the acts of:

obtaining data corresponding to media content from a handheld device, the data representing picture elements of video or representing audible portions of an audio signal;

computing attributes of the data using a processor, said act of computing utilizes the processor to operate on the data;

using computed attributes of the data to identify the media content or to identify metadata associated with the media content;

obtaining metadata associated with the media content; and

providing the metadata to the handheld device from a network resource.

38. A programmed computing device comprising the one or more computer readable media of claim **37**.

39. A method comprising:

obtaining data corresponding to media content from a process in a handheld device, the data representing picture elements of video or representing audible portions of an audio signal;

computing attributes of the data using a processor, said act of computing utilizes the processor to operate on the data;

using computed attributes of the data to identify the media content or to identify metadata associated with the media content;

obtaining metadata associated with the media content from a network resource;

providing the metadata for output or display.

40. The method of claim **39** wherein the multi-purpose electronic processor is located remotely with respect to the handheld device.

41. The method of claim **39** wherein the handheld device comprises the multi-purpose electronic processor.

42. A programmed computing device storing instructions in memory, said instructions are executable by said programmed computing device to perform the acts of claim **39**.

43. A computer readable media comprising instructions stored thereon to cause a multi-purpose electronic processor to perform the acts of claim **39**.

44. The method of claim **39** in which said act of using computed attributes of the data to identify the media content or to identify metadata associated with the media content identifies metadata associated with the media content.

45. The method of claim **39** in which said act of using computed attributes of the data to identify the media content or to identify metadata associated with the media content identifies the media content.

46. The method of claim **39** in which the data represents the audible portions of an audio signal or represents picture elements of video.

47. The method of claim **39** in which the handheld device comprises a cell phone.

48. One or more computer readable media comprising instructions stored thereon that when executed by one or more processors perform the acts of:

obtaining data corresponding to media content from a process in a handheld device, the data representing audible portions of an audio signal;

computing attributes of the data using a processor, said act of computing utilizes the processor configured to operate on the data;

using said computed attributes of the data to identify the media content or to identify metadata associated with the media content;

obtaining metadata associated with the media content from a network resource;

providing the metadata for output or display.

49. A programmed computing device comprising the one or more computer readable media of claim **48**.

50. The one or more computer readable media of claim **48** in which the handheld device comprises a cell phone.

51. One or more computer readable media comprising instructions stored thereon that when executed by one or more processors perform the acts of:

obtaining data corresponding to media content from a process in a handheld device, the data representing picture elements of video or representing audible portions of an audio signal;

computing attributes of the data using a processor, said act of computing utilizes the processor to operate on the data;

using computed attributes of the data to identify the media content or to identify metadata associated with the media content;

obtaining metadata associated with the media content from a network resource;

providing the metadata for output or display.

52. The system of claim **51** in which the multi-purpose processor is operating to perform at least one of the functions recited therein.

53. The system of claim **51** in which the handheld device comprises a cell phone.

54. A method comprising:

obtaining data corresponding to media content from a process in a handheld device, the data representing audible portions of an audio signal;

using an electronic processor configured to compute attributes of the data;

using attributes of the data computed by the electronic processor to identify the media content or to identify metadata associated with the media content;

obtaining metadata associated with the media content from a network resource;

providing the metadata for output or display.

55. A programmed computing device storing instructions in memory, said instructions are executable by said programmed computing device to perform the acts of claim **54**.

56. A computer readable media comprising instructions stored thereon to cause an electronic processor to perform the acts of claim **54**.

57. The method of claim **54** in which said act of using attributes of the data computed by the electronic processor to identify the media content or to identify metadata associated with the media content identifies metadata associated with the media content.

58. The method of claim **54** in which said act of using attributes of the data computed by the electronic processor to identify the media content or to identify metadata associated with the media content identifies the media content.

59. The method of claim **54** in which the data represents the audible portions of an audio signal or represents picture elements of video.

60. The method of claim **54** in which the handheld device comprises a cell phone.

61. A system comprising:

an input for obtaining data corresponding to media content from a handheld device, the data representing audible portions of an audio signal;

an electronic processor configured to: i) compute attributes of the data; ii) using at least said computed attributes of the media content, identify the media content or identify metadata associated with the media content; iii) obtain metadata associated with the media content from a network resource; and iv) provide the metadata for output or display.

62. The system of claim **61** wherein the electronic processor is operating to perform at least one function recited therein.

63. The system of claim **61** in which the handheld device comprises a cell phone.

64. A system comprising:

means for obtaining data corresponding to media content from a handheld device, the data representing picture elements of video or representing audible portions of an audio signal;

means for computing attributes of the data;

means for using at least computed attributes of the media content to identify the media content or to identify metadata associated with the media content;

means for obtaining metadata associated with the media content from a network resource; and

means for providing the metadata for output or display.

* * * * *